This is a digital copy of a book that was preserved for generations on library shelves before it was carefully scanned by Google as part of a project to make the world's books discoverable online.

It has survived long enough for the copyright to expire and the book to enter the public domain. A public domain book is one that was never subject to copyright or whose legal copyright term has expired. Whether a book is in the public domain may vary country to country. Public domain books are our gateways to the past, representing a wealth of history, culture and knowledge that's often difficult to discover.

Marks, notations and other marginalia present in the original volume will appear in this file - a reminder of this book's long journey from the publisher to a library and finally to you.

Usage guidelines

Google is proud to partner with libraries to digitize public domain materials and make them widely accessible. Public domain books belong to the public and we are merely their custodians. Nevertheless, this work is expensive, so in order to keep providing this resource, we have taken steps to prevent abuse by commercial parties, including placing technical restrictions on automated querying.

We also ask that you:

+ *Make non-commercial use of the files* We designed Google Book Search for use by individuals, and we request that you use these files for personal, non-commercial purposes.

+ *Refrain from automated querying* Do not send automated queries of any sort to Google's system: If you are conducting research on machine translation, optical character recognition or other areas where access to a large amount of text is helpful, please contact us. We encourage the use of public domain materials for these purposes and may be able to help.

+ *Maintain attribution* The Google "watermark" you see on each file is essential for informing people about this project and helping them find additional materials through Google Book Search. Please do not remove it.

+ *Keep it legal* Whatever your use, remember that you are responsible for ensuring that what you are doing is legal. Do not assume that just because we believe a book is in the public domain for users in the United States, that the work is also in the public domain for users in other countries. Whether a book is still in copyright varies from country to country, and we can't offer guidance on whether any specific use of any specific book is allowed. Please do not assume that a book's appearance in Google Book Search means it can be used in any manner anywhere in the world. Copyright infringement liability can be quite severe.

About Google Book Search

Google's mission is to organize the world's information and to make it universally accessible and useful. Google Book Search helps readers discover the world's books while helping authors and publishers reach new audiences. You can search through the full text of this book on the web at `http://books.google.com/`

BF
1
.A67
No. 197

The Effects of the Relative Difficulty of Mental Tests on Patterns of Mental Organization

B 818,116

BY
MAX HERTZMAN

Submitted in partial fulfillment of the requirements for the degree
of Doctor of Philosophy, in the Faculty of Philosophy,
Columbia University

ARCHIVES OF PSYCHOLOGY
R. S. WOODWORTH, EDITOR
No. 197

NEW YORK
April, 1936

ARTES SCIENTIA VERITAS 1817
LIBRARY OF THE
UNIVERSITY OF MICHIGAN

RECEIVED IN EXCHANGE
FROM
Columbia University
Library

The Effects of the Relative Difficulty of Mental Tests on Patterns of Mental Organization

BY
MAX HERTZMAN

Submitted in partial fulfillment of the requirements for the degree
of Doctor of Philosophy, in the Faculty of Philosophy,
Columbia University

ARCHIVES OF PSYCHOLOGY
R. S. WOODWORTH, EDITOR
No. 197

NEW YORK
April, 1936

LF
1
.A67

Exchange
Columbia University
Library

JUN 17 1937

TABLE OF CONTENTS

PAGE

CHAPTER I. PROBLEM AND LITERATURE 5
1. The problem 5
2. The literature 7

CHAPTER II. PROCEDURE 13
1. The subjects 13
2. The tests 13
3. The testing procedure 14
4. The division into levels 14
5. The factor of speed 15
6. The problem of reliability 17

CHAPTER III. RESULTS 20
1. Evaluation of the variables 20
2. Analysis of the correlations 23
3. Analysis through the use of mean tetrads 35
4. Analysis by the use of triads and by an iterative factor method 46
5. Comparison of the derived variables with the 'whole' test variables 59

CHAPTER IV. CONCLUSIONS AND INTERPRETATIONS 64

BIBLIOGRAPHY 67

CHAPTER I

PROBLEM AND LITERATURE

1. THE PROBLEM[1, 2]

The past few years have witnessed a tremendous development in the use of the correlation coefficient for the analysis of psychological data and the multiplication of techniques which attempt analysis beyond the correlation coefficient itself. These latter methods, starting with the matrices of correlation coefficients as a base, have attempted to organize the tests between which the correlations are obtained in terms of the sums of independent mathematical variables.

However, no matter whether we treat the correlations themselves or the patterns derived from them as fundamental in our analysis of the relations between and among test variables, we find that in either case a particular test variable is treated as a unit. The correlation coefficient cannot indicate whether the items which make up the test variables in question contribute equally or unequally to the degree of concomitant variation of which the coefficient is a measure. It cannot tell us whether there exist opposed relational tendencies within different aspects of the variables themselves which, though they possibly measure heterogeneous functions must nevertheless be treated by the technique of correlation as homogeneous units. In the analysis by factors, there is an attempt to break up the seemingly rough measure of variance that the correlation coefficient implies into several uncorrelated components. This, however, hardly amounts to saying more than that while we have measured the relative mathematical dependence of two variables on each other by one correlation coefficient, we may now consider this relationship as expressed by the sum of several correlation coefficients which add up to the original one.[3] In the latter method, the various uncorrelated tendencies which the factors themselves are supposed to express are, it is to be noted, tendencies which, while they are not related to each other, are nevertheless thought of as being representative of the entire variable.

---

[1] This is one of a series of studies under the general direction of Professor H. E. Garrett.

[2] The writer wishes to express his sincere appreciation and gratitude to Dr. G. M. Smith for the latter's kindness in making available to the writer the original data that he used in his study (27).

[3] This method does have the advantage of expressing the various relationships among the test variables in terms of a relatively limited number of mathematical variables rather than in terms of $n(n-1)/2$ individual constants.

It is very possible that the items which compose a particular mental test, while presumably measuring the same function, may do so to widely different degrees. Some of them may bear so little a relation to what others are measuring that items from tests defined by their originators as measuring a different function may actually bear a stronger relation to them. It is difficult, when devising a test, to make a critical analysis in order to obtain internal homogeneity. The necessity for intercorrelating the items makes it rare that one will select items for an entire battery in such a way. The ordinary methods of test construction are, however, of such a nature as to introduce arbitrarily a source of heterogeneity. It is usual to introduce items of differential difficulty so that they may be arranged from the easiest to the most difficult. This is, no doubt, an important technique for establishing rapport with the subject taking the test, but it does introduce a heterogeneous factor as regards one important reference point, namely, the level of difficulty. To find whether or not this type of heterogeneity introduces a functional heterogeneity, in the sense that easy items show different relationships in similar set ups than the difficult ones, is one of the main purposes of this study.

The results of previous studies on the interrelationships among test variables show in many cases strong evidence for different degrees of relationship among different types of variables. Thus, Schneck (26) shows that functions measured by different verbal tests are likely to be more closely related to each other than they are related to functions measured by numerical tests. Smith (27) has demonstrated the same fact and, in addition, similar ones for the relations between tests measuring ability to deal with spatial relations and verbal ability, and tests of 'spatial' ability and numerical ability. It becomes a pertinent question to ask whether these relationships will hold at all, or if they do, whether they will hold to the same degree when the easy and the difficult material which together comprise the tests between which the previously mentioned relationships have been found are treated as separate variables.

Consequently the present problem took the form of studying the various kinds of relationships which obtain between variables of different material, which in a previous analysis had been shown to be functionally different, when each of the original variables was divided into two variables, one consisting of the scores made on the fifty-percent easiest items, and the other of the scores made on the remaining items.

In a sense this problem may then be thought of as being a study of some particular set of conditions which influence the 'constancy' of the mathematical patterns derived to describe mental organization. Previous results, as will be shown later, have indicated that the structure of these patterns may change with the age of the group upon which they have been determined. This study will attempt to find whether the patterns may be affected by the heterogeneity introduced through having the items which make up the test variables vary in difficulty.

The method of defining the variables under consideration here makes it possible to study the effects of heterogeneity in an additional manner to that mentioned above. Since the extent of the relationship existing between the relatively easy and relatively difficult aspects of the same variable can be determined, we shall be able to see whether there are functional relations between different variables which are greater than the two different parts of what had previously been treated as a single homogeneous function. These analyses may shed some light on the conditions other than similarity of material which determine the degree of the functional relation between different test variables. The division of the original tests into easy and difficult variable will also enable us to study the effects of the level of difficulty on the interrelationships among tests of the same material.

Briefly summarized, the problems of this study may then be stated as being: (1) to find whether the level of difficulty of a test will affect its relationship with tests of different material; (2) to find whether the level of difficulty of a test will affect its relationship with other tests of the same material; (3) to study the implications of the results obtained from working out the first two problems for the general problem of mental organization.

2. THE LITERATURE

In Spearman's 'Abilities of Man', (30, pp. 213 ff), it is suggested that the degree of difficulty of the items in mental tests may play an important role in the magnitude of the correlations that are found among these tests. On the basis of a study by one of his students, Hamid (16), he (and the latter) conclude that material of medium difficulty is probably most suitable for obtaining the highest intercorrelations. Spearman points out, however, that the effect of difficulty will vary, depending on the general level of the group on which the data is obtained. "With less capable subjects, the advantage will perhaps shift round from the medium to the smaller degrees of difficulty; with more capable subjects, conversely, the advantage would go over to the greater degree" (30, 213). We must remember, however, that the terms 'easy', 'medium', and 'difficult' are in reality relative to some standardization group. If Spearman's statement quoted above is to have any consistent meaning, the reference point must be some group of medium ability. For groups with higher and lower abilities, the more difficult and less difficult items will become, respectively, the items of medium difficulty. This is consistent with the general view of mental test constructors that too easy or too difficult items have very little discriminating capacities. It sheds no light on the boundaries to be set in order to determine medium difficulty, which must always be determined for the particular set of conditions under consideration. It should be noted that Spearman's statement should not be taken to mean that in any distribution of the relative difficulty of the items of a test for a given group those items of medium difficulty are the most discriminating. The scale of difficulty might not be exhaustive so that for different groups of subjects there might exist items not included on a given test which would be far more discriminating than any that are a part of it.

The interests of Thorndike in a test like the CAVD, supposedly measuring different levels of difficulty, have resulted in some studies by himself and his students on the relation between easy and difficult material. In "The Measurement of Intelligence", Thorndike presents evidence indicating that the more remote from each other different levels of the CAVD are, the smaller their correlations will be. Thus, he presents correlations (33, 558, Table 154) between levels N and O of .68, but between N and Q of .47; between O and P, .65, while between O and Q there is a correlation of .43. In another table (p. 560) we find that the correlation between I and J is equal to .59, while that between I and M is .49. Similar results may be found for various other combinations of levels, even at the very low ones. Using the levels A to D on a group of 98 imbeciles, Thorndike finds a correlation between A and B of .78 against .52 for A and D. For B and C, a correlation of .81 as compared to one of .59 between B and D is reported. (p. 557.)

It is obvious that the majority of the correlations that are reported here are far from unity. It is therefore quite remarkable to find in a later table (Table 160, p. 562) that the average correlation between levels one step removed is .968 and between levels two steps removed .980. This is accomplished by the most unrestrained use of the correction for attenuation and a correction applied to nullify the effects of a 'limited' range. From the corrected coefficients Thorndike proceeds to draw the conclusion "that the correlations between composites far apart in difficulty are little, and possibly not at all, lower than the correlations between neighboring composites," (33, 564) a conclusion which is not warranted by the uncorrected coefficients. The tremendous rise in the size of the correlation coefficient from the moderate ones we have taken from Thorndike to his final corrected ones is due in part to his method of ascertaining the reliability coefficients. This latter coefficient is obtained from correlations of tests with only twenty items each. The small number of items sampled, alone, is influential in depressing the reliability coefficients thus obtained. But it should also be remembered that the twenty items contain four different materials thus reducing the homogeneity of the test function.[4] While this diversity of material will of course affect the correlations between different levels also, the effect of a lowered reliability coefficient, if corrections for attenuation are used, is more thoroughgoing and can result in large gains for the corrected coefficient. Even with the attenuation correction, correlations were in most cases raised only to .85 or .90 and corrections for a restricted range were able to account for the rest of the gain. In all fairness, we do not think that there can be any legitimate use of the latter correction for Thorndike's data. In making that correction, in effect, a hypothetical range and distribution are substituted for the obtained ones. Since the levels used by Thorndike are each supposed to measure a definite degree of difficulty, and restricted ranges are obtained, it is very possible that the corrected distribution could be obtained empirically only if material heterogeneous with respect to level were introduced. But this would defeat the very purpose of the analysis. We feel that Thorndike's data on the CAVD indicates clearly that remoteness of level of difficulty does depress the correlation between two tests.

---

[4] Garrett (13, p. 106) presents intercorrelations between different parts of the CAVD, that is, between the four different tests, for levels M, N, O, P, and Q, obtained on 313 Brooklyn College girls. These correlations range from .58 to .21 indicating a diversity of relationships among these tests.

Peatman (23) has shown that the tests in the Thorndike Entrance Examination for high-school graduates which have some similarity to the CAVD may be accounted for in terms of a common factor plus the necessary specifics. It should be noted, however, that this conclusion is arrived at through the use of tetrad analysis which allows considerable leeway in the variability of the correlation coefficients. In a set up which actually satisfied the tetrad difference criterion, there were correlations ranging from .49 to .15 (P. 34, Table VIII).

From a study by Tilton (37), utilizing material other than the CAVD, Thorndike concludes that the causative factors of higher abilities and lower abilities are not essentially different from each other, (33, 425). In the latter study, sentence completion tests, arithmetical problems tests, and analogies tests were used as measures of the higher processes. For the lower or "purely associative" processes, Tilton used vocabulary tests, routine and informational arithmetic tests, and information tests. The subjects were 250 boys who were just graduated from the eighth grade. Thorndike takes the purely associative processes to be "lower" processes, but neither he nor Tilton present an adequate standard for this designation. Certainly one would hesitate to decide that the vocabulary test is measuring a different level of ability than the sentence completion test without taking as a criterion the actual difficulty of the items of the two tests as determined by the success of the group with them. Actually, Tilton, (37, 20) presents a distribution of the frequency of the number of subjects passing a varying percent of items for each test. The data are grouped, and using the midpoint of the group as representative of the total frequency within it, the present writer calculated the percent of the items that were passed correctly for both the 'higher' and the 'lower' tests. For the former 76.24% of the total items were passed, whereas for the latter, the percentage was 68.44. This difference is not very great, but whatever difference in difficulty they indicate, it is in the direction of showing that the items testing the lower processes were the more difficult ones. Tilton finds that the average correlation of higher with lower tests is .558, of higher with higher .544, and of lower with lower .571. In view of the relative equality in difficulty of the items in the two sets of tests, these results are not surprising. Thorndike, taking the correlations presented by Tilton, applies the attenuation formula in such a way as to conclude that there is a perfect relationship between what is common to any two association tests and any two power tests. This is done by assuming that the correlation between any two association tests may be treated as a reliability coefficient and likewise as regards the corresponding measure between any two power tests. Needless to say, these correlations are lower than those which ordinarily would be considered true reliability coefficients and consequently they raise the corrected figure more than a usual reliability coefficient would. In view of the relative uniformity of the three groups of correlations concerned in this particular case of attenuation correction, it is easy to see how the corrected figure will become equal to, or about, unity.[5]

In a similar fashion, Clark (5), using Thorndike's formula finds correlations between what he calls 'general level' and 'intelligence' which are very close to unity. Intelligence, apparently, is measured in his study by the Stanford-Binet test, the Otis Self-Administering Group test, and the Terman Group test; while general level is measured by an arithmetical problems test and a sentence completion test. The correlation between the arithmetical problems test and the sentence completion test is used in the attenuation formula as if it were a reliability coefficient. Its value which is equal to .55 explains why the corrected coefficients loom so large. Thus, Clark transforms two correlations of .80 and .66 between the arithmetical problems test and the Terman and the sentence completion and the Terman, respectively, into a correlation of approximately one (1.03). It is difficult for us to see what is gained by such a technique except additional confusion in an already highly complicated field.

We feel that the work of Thorndike and his students has important bearing on many problems of mental organization, but that many of the results are distorted by a too generous use of corrections. The attenuation formula is given a dangerous elasticity and it is on the basis of the 'corrected' results rather than on the original ones that many of the major, and unfortunately, better known conclusions of that group are drawn.

In our problem, we wish to see whether different degrees of difficulty in the test items will affect the patterns of mental organization. There has been previous work indicating that the relationships among different tests have not been stable ones in the sense that they remain the same with the development of the subjects under consideration. Indirect evidence with regard to this point may be obtained by comparing the study of Schneck (26) with that of Schiller (25). The former finds an average correlation between four verbal and four numerical tests of .1441 obtained on 210 City College sophomores. The latter, using 189 third and fourth grade boys, finds the average correlation between four verbal and three numerical tests to be .6291. (Similar results were obtained with the girls in this study.)

---

[5] If a correlation equal to a given constant is divided by the product of the square roots of two other constants which are about equal in magnitude to the first one and to each other, the result must be equal to unity or a number very near to it.

Garrett, Bryan and Perl (15) have recently made a more systematic investigation of the age factor in mental organization. They used as subjects three groups of children of ages 9, 12, and 15, respectively. The tests used included six measuring memory performances; and one each of vocabulary, arithmetic, spatial relations, and speed of response. It was found that the average correlations decreased with age. When factor analysis was used, the weights of the largest factor were also found to decrease with the increase of age.

The most direct approach to the problem of the relation of patterns of mental organization to age, is that of Asch (3). He was able to retest some of the original subjects of Dr. Schiller three years after the first testing; and since the original scores were made available, he was able to compare the two sets of relationships.[6] At the first testing, the subjects averaged nine years of age and at the second, twelve years. The results show a reduction of the correlations with the development of the children. The average correlation of the boys dropped from .56 to .41, and for the girls from .59 to .51. When correlations between verbal and numerical abilities are compared, the drop is found to be greater than when the tests compared are homogeneous with regard to material.

These preceding studies show quite clearly that one need not expect constancy in the relationships between different test variables. Development with age seems to bring about a certain degree of specialization which is shown by the effects on correlations between tests of different material. Our study attempts from another angle to discover factors which may alter the relationships among test variables.

---

[6] Of the original subjects he was able to obtain 79 of the boys and 82 of the girls.

CHAPTER II

THE PROCEDURE

The data used in this study were obtained from seven of the original fourteen tests that had been given by Dr. Smith, as reported in his monograph (27). We do not think it necessary to repeat it any great detail his excellent report on the characteristics of the subjects, the nature of the tests, and the testing procedure all of which may be found in his monograph (27, 13–20).

1. THE SUBJECTS

The 186 subjects constituted a very homogeneous group with reference to age, sex, race, socio-economic status, and educational background. This is desirable, as homogeneity has been shown by Kelley (21) and Garrett and Anastasi (14) to be of prime importance in eliminating spurious correlation. The subjects were all of the male sex; the average age was 19 years and 5 months with a rather small standard deviation of 16 months. The subjects who were students at the College of the City of New York were not widely scattered through the possible classes, as 80% were either in the upper sophomore or junior class.

2. THE TESTS

Seven of the original fourteen tests were used, four of them being numerical tests and three spatial tests. These two material groups were chosen rather than other possible pairs as the original work had shown that they were more closely related to each other than other pairs, and it was hoped that they would consequently be more sensitive to the influence of the new set of conditions introduced by our analysis. The Minnesota Paper Form Board Test was dropped out of the original spatial battery as it was a 'speed' test rather than a 'power' test. Even though it is a speed test, it had fairly high correlations with the other spatial tests (average correlation was equal to .54) and this may have justified its inclusion. But as a speed test, it was in such form as to make the concept of varying difficulty of items almost a meaningless one. Very few items, if attempted, were done incorrectly. On the contrary, the final score was almost a direct function of the number of items that the subject had time to attempt. Consequently a division into easy and difficult items on the basis of the percentage of correct answers for each item would not have been feasible.

The remaining spatial tests used by Smith and the ones utilized in this study were the spatial analogies test, the spatial generalizations test and the modified Kelley spatial test. Illustrations of items from some of these tests and the numerical ones may be found in Smith's paper (27, 16–7). The last mentioned spatial test was modified from one described by Kelley (21). It was changed so as to be difficult enough for college students. This test, as a matter of fact, turned out to be one of the most difficult of the entire battery. One of the chief tasks required of the subject taking it was "to determine (in imagination) which of certain elementary shapes need to be turned over (not merely rotated in the same plane) in order to fit the larger given figures." (27, 18.) All of the numerical tests used in Smith's study were retained. They were the number series completion test the numerical analogies, the numerical generalizations, and the arithmetical reasoning tests. They were of sufficient difficulty as shown by the results to test the college level ability.

3. THE TESTING PROCEDURE

The procedure used in testing is described fully by Smith (27, 18–19). It was very likely that rapport had been fully established in most of the cases. The subjects were 'test-wise', having had considerable experience with psychological tests before. According to Smith, "There was every indication that the students, with very few exceptions, accepted the tests as a challenge to their intellectual powers and put forth their best efforts." (27,19.)

4. THE DIVISION INTO LEVELS

In dividing the tests used into only two levels of difficulty, a certain amount of arbitrariness, of course, was involved. This, however, was necessitated by the limited number of items to be found in a particular test. The modified Kelley spatial with the largest number of items gave tests of only 23 items each when it was divided into two tests. The number series completion test which had the fewest items (30) was reduced to two tests of 15 items each. Even 23 items, the most possible for a single variable under our conditions, were not as many as one would ideally desire. Further division, even if it were only into three levels, would have given us tests with as few as 10 items, while the longest tests would have had only 16. Under such circumstances, making more than two levels for each test would have been impractical.

The items for each test were tallied individually with reference to the total number of errors made by the group on each one of them. In this way, an index of the difficulty of an item was obtained, difficulty being defined here purely in terms of the relative success of the group. The items were arranged according to the percentage of errors made on each one. Each test was subdivided into two variables, called arbitrarily an easy and a difficult variable. The easy variable consisted of the 50% easiest items in terms of the percentage of errors for each item, and the difficult variable, of the remaining items. There was no sharp line drawn between the items included in the two kinds of variables derived from each test, as the percentage of errors for the items in each test varied with fair regularity. There was no particular point in excluding items around the dividing line between the two sets of variable merely to accentuate the difference between the two groups. As results presented later will show, the difference between the average level of the two groups was in the case of each test a rather wide one. The elimination of borderline items would only have reduced the size of our already rather small derived tests.

5. THE FACTOR OF SPEED

As already suggested in the discussion of the elimination of the Minnesota Paper Form Board test from our analysis, a speed test in which one could not distinguish the relative difficulty of the items by the results on the items attempted, would be very undesirable for our purposes. In the tests as originally given, while the time allowed had to be limited because of reasons beyond the experimenter's control, enough time was allowed to make the tests, for the most part, power tests. In addition, those tests which when scored showed that many subjects had not finished, were cut down, a variable number of the terminal items (varying for each test) being eliminated from the final score. With this arrangement, the percentage of subjects finishing each item ranged among the tests from 74 to 88 percent, with a median of 80% (27, 20).

This introduces, though in a very limited fashion, a premium on speed. One should note, however, that the slight speed factor that is introduced here is not of the same kind as that which one would find, for example, in a test where few errors are made, and where consequently the individual items are not very difficult. It has been customary to assume that speed in 'intelligence' tests should be measured at a very low level of difficulty, and then it is easy to show that such speed has only a slightly relationship to the ability exhibited on power tests. But speed of performance is very likely linked with the difficulty of material upon which it is measured, and the relation of speed to general test ability cannot be assumed to be the same on all levels of difficulty. Dubois (7), in following the first concept mentioned above, utilized tests as a measure of speed that were so easy for his college student subjects that virtually no errors were made on them. The average correlation of four of his speed tests with a standard arithmetical reasoning test was only .1817 and with a vocabulary test, .2721. The average intercorrelations of the speed tests with each other was considerably higher, being equal to .5764 (7, 27). In our tests, if speed were a factor, it could not appear in a form wherein Dubois' results were relevant. By utilizing a group previous to the main experimental group, Dr. Smith had arrived at a tentative order of difficulty of his test items, the final tests being arranged in this order. As a result we find that virtually none of the easy items are in the last quarter of any test, and even fewer in the last third. It is the last quarter from which almost all of the omitted items come. We can then hardly expect that speed could have influenced the scores on the easy items which now comprise the so-called easy variables, and thus be a factor in introducing spurious correlations.

If speed is effective, it will show its effects on the difficult variables. But it should be remembered that Dubois' results and others like his cannot be made to apply to our difficult variables. The situation is really analogous to one in which we have a difficult test given with a time limit which does not allow all the subjects to complete the test. We wish to find if this seriously alters the relationship obtained when the subjects have an unlimited time for the test. A study by Freeman (10) is of great interest with respect to this point. He gave several standard intelligence tests with a limited time and other forms of these tests with unlimited time. With the exception of one of these tests, the correlations are remarkably high. With the National Intelligence Test, the correlation is .83; when the Dearborn Group Test is used, the correlation is .88; and it is .93 with the Terman Group Test. The Otis Advanced Group Test correlation is the only one that is low, that correlation being .58. Freeman offers no explanation for this exception, nor are the conditions presented in sufficient detail for us to hazard one. Despite the correlation of .58, the average of these correlations is still very high, being .81. If Fisher's 'z' were used (9), the high correlations would be weighted more than the low ones and our average would be even higher. Similar results have been obtained by correlating the Army Alpha Test given with a normal time limit and with double time. A correlation of .967 on 510 soldiers is reported (41, 417). While this correlation is unquestionably influenced by the wide range of ability in the group tested as compared to the rather restricted range of our group, it is nevertheless very suggestive. For the majority of the group, the test was a power test and increasing the time did not change the relative rankings to any great extent.

In view of the facts, (1) that relatively few of the items that were used in the final tabulation were not attempted; (2) that those not attempted fell largely in the difficult half of the test; and also in view of (3) the high correlations reported above from the work of Freeman and the results on the Army Alpha, we feel jusified in concluding that our correlational results do not differ very widely from what would have been obtained should it have been possible to allow unlimited time for the original tests.

6. THE PROBLEM OF RELIABILITY

Reliability is often defined as the consistency of a test or questionnaire as a measuring device. But actually, the reliability coefficients are obtained in various ways; such as by correlating odd and even items, correlating 'comparable' forms of the same test, and by correlating the results of a test with the results of a retest at a later date. As has been correctly pointed out by Willoughby (39), and Peatman and Greenspan (24), these different methods of obtaining the coefficient imply that the definition of reliability is an operational one, and the meaning of the reliability coefficient differs, depending on which method of obtaining it has been used.

It becomes pertinent in a correlational study such as this one to ask what role reliability coefficients may play in interpreting the data; and what relationships the reliability coefficients that are used (or in this case, the only ones which can be obtained) imply. In correlation studies, fairly large sized reliability coefficients are taken as an assurance that the function measured is relatively stable. The square root of the reliability coefficient may be used as a practical gauge in judging the probable upper limit of any particular correlation (34). In this study, the tests could not be presented for retest for many reasons. There were so many tests that it would have been unreasonable to require the subjects to take them again, and the amount of time required, even if the subjects had been

ARTES 1817 SCIENTIA
VERITAS
LIBRARY OF THE
UNIVERSITY OF MICHIGAN

RECEIVED IN EXCHANGE
FROM
Columbia University
Library

The Effects of the Relative Difficulty of Mental Tests on Patterns of Mental Organization

BY
MAX HERTZMAN

Submitted in partial fulfillment of the requirements for the degree
of Doctor of Philosophy, in the Faculty of Philosophy,
Columbia University

ARCHIVES OF PSYCHOLOGY
R. S. WOODWORTH, EDITOR
No. 197

NEW YORK
April, 1936

lation coefficient in psychological studies. The coefficient was originally derived on the assumptions of the normality of the distributions in question, and on the assumption of the existence of linear regression. They point out that few investigators take these facts into consideration. We certainly cannot expect that our skewed curves will show linear regression, and we would be totally unjustified if we used correlations based on them for the purpose of predicting individual scores in a regression equation. Our concern, however, is not with this type of prediction, but with the effect of skewness on the value of our correlation coefficient. Smith, when faced with a similar difficulty, made his most extremely positively skewed and his most extremely negatively skewed curves artificially symmetrical. The effect was to change the correlation of .476 to .470, a difference of .006 (27, 23). This would indicate that skewness in this set-up did not greatly affect the value of the correlation coefficient.

With our own correlations, we applied another test. Since we had divided our variables into easy and difficult ones we could use the formula for the correlation of sums (28) to predict the original correlations of the whole variables that had been reported by Smith. Thus, the correlation between any two variables, x and y could be thought of as being represented by $r_{(E_x+D_x)(E_y+D_y)}$. Using equal weights for the correlations we had obtained (See Table III), the original correlations were predicted. Table II presents the absolute average difference between the obtained and the predicted correlations, not only for all the correlations under consideration, but also for the correlations of each variable with each of the others. The average for all variables differs only by .0369 from the obtained correlations, with a critical ratio of 1.48. For individual variables, the deviations from the originals are also low, the greatest difference being .0581 for the arithmetical problems. When we remember that the formula under the best of conditions will give us only approximations, these predicted results seem to indicate that the skewness of our sub-test variables have not markedly affected our correlations.

Our measure of skewness is very sensitive and may be affected by a few extreme deviations. That this is probably so, is shown by the values of our measures of kurtosis which, in part, help to explain the excellent prediction just reported. A perfectly mesokurtic distribution has a kurtosis equal to .263. With our standard error of kurtosis being equal to .020, we can observe from the table that many of our variables fall very close to the mesokurtic ideal, while none

TABLE 2

EVALUATION OF THE DIFFERENCES BETWEEN OBTAINED AND 'PREDICTED' VALUES OF THE CORRELATION COEFFICIENT*

Test Variables	*Num. Analog.*	*Num. Gen.*	*Arith. Prob.*	*Num. Series Compl.*	*Spatial Analog.*	*Spatial Gen.*	*M.K. Spatial*	*Average***
Absolute diff. between obtained and predicted r's	.0292	.0378	.0581	.0452	.0258	.0308	.0313	.0369
D/σ_D	1.16	3.00	2.55	1.54	2.21	1.15	1.25	1.48

* Test variables refer to the original undivided tests. In general, the predicted values are lower than the obtained ones.

** The number of correlations in determining the average was 21. Each individual variable was involved in 6 correlations.

fall beyond three standard errors of it. Variables 1, 3, 5, 6, 7, II, and V are within one standard error of perfect kurtosis; while the remaining variables are within two standard errors.

2. ANALYSIS OF THE CORRELATIONS

a) Relation among combinations of easy and difficult variables.

Table III presents the intercorrelations of the fourteen variables of the study.[10] Because of the nature of the variables under consideration these correlations may be grouped in various ways. If we consider only the easy-difficult dichotomy, four groups of correlations may be formed from the original table. They consist of (1) intercorrelations among all the easy variables; (2) intercorrelations among all the difficult variables; (3) cross correlations involving the various easy-difficult correlations, with the exception of easy-difficult correlations derived from what were originally the same variables; and (4) correlations between easy and difficult halves of what was originally the same variable. These groups may be symbolized, respectively, as E–E; D–D; E–D; and E–D$_{(s)}$. Our immediate concern is with the first three kinds of correlations.

---

[10] These correlations seem unusually small as compared to correlations ordinarily reported in the literature of mental tests. It should be remembered, however, that the correlations are based on tests much shorter than those usually used in these studies. The curtailed size has the effect of depressing the correlation coefficient. In addition, the great homogeneity of this group and its relatively high general level of ability is also a factor in making for low intercorrelations.

TABLE III

Intercorrelations of the 14 Variables*

	Variables	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*I*	*II*	*III*	*IV*	*V*	*VI*	*VII*
1.	Easy Num. Analogies		.2792	.2558	.1678	.3211	.2280	.2701	.4951	.3393	.2156	.3484	.2239	.2203	.2726
2.	" " Generalizations	.287		.2972	.2639	.3406	.3016	.2667	.3257	.5799	.2439	.2839	.3015	.2581	.2094
3.	" Arith. Problems	.261	.307		.2802	.4797	.3367	.4584	.3050	.4054	.5504	.4713	.3672	.2584	.2552
4.	" Number Series Completions	.169	.270	.288		.2968	.2730	.2548	.2295	.2582	.2439	.3433	.2615	.2810	.1497
5.	" Spat. Analogies	.333	.355	.522	.306		.3776	.4757	.3797	.3840	.4499	.4585	.4279	.2260	.3121
6.	" " Generalizations	.232	.311	.350	.280	.397		.3859	.2277	.4082	.2013	.3744	.3428	.4840	.3511
7.	" M.K. Spatial	.277	.273	.495	.261	.517	.407		.3011	.3589	.3478	.4603	.4385	.3102	.5814
I.	Diff. Num. Analogies	.543	.338	.315	.234	.400	.232	.311		.4154	.3142	.3587	.4537	.3315	.3712
II.	" " Generalizations	.353	.661	.430	.262	.405	.433	.376	.442		.4135	.3952	.3962	.3838	.2873
III.	" Arith. Problems	.219	.249	.619	.249	.485	.204	.363	.325	.440		.4270	.3838	.2684	.3159
IV.	" Number Series Completions	.363	.292	.512	.352	.496	.393	.493	.375	.418	.456		.4061	.4086	.3717
V.	" Spat. Analogies	.228	.311	.395	.263	.458	.358	.470	.481	.419	.404	.431		.3997	.4546
VI.	" " Generalizations	.224	.265	.262	.289	.230	.528	.321	.344	.404	.275	.434	.423		.3999
VII.	" M.K. Spatial	.280	.212	.261	.151	.323	.367	.664	.390	.296	.327	.390	.490	.416	

* Original correlations (Pearson) above the diagonal; the corresponding Fisher's 'z' coefficients below the diagonal. The Pearson correlations were calculated by the Columbia University Statistical Bureau.

In Table IV we have the means,[11] standard deviations, and coefficients of variation of the first three types of correlation. It is interesting to note that there is virtually no difference between the average of the E–E group intercorrelations and those of the E–D. If the E–D correlations are split up into two groups, one containing correlations whose variables consist of the same kind of material, and the other set containing correlations between variables of different materials, the situation is altered very little. The average of the former group is .327 and of the latter .320. The difference between the D–D group and each of the first two, while apparently rather small, is reliable. The ratio of the difference to the sigma of the difference when D–D and E–E are considered is 3.04 and when E–D is considered with D–D, the ratio is 3.30. Reliable differences between E–D correlations and D–D correlations are also obtained when E–D correlations that are similar in material and those that are different in material are considered separately.

TABLE IV

EVALUATION OF VARIOUS COMBINATIONS INVOLVING EASY AND DIFFICULT CORRELATIONS (z'S)

Combinations of Correlations	*Mean*	*S.D.*	*C.V.**	*Evaluation of Selected Differences*	
				Combinations Evaluated	D/σ_D
(a) E–E	.329	.091	27.66	(a) and (b)	3.04
(b) D–D	.399	.056	14.04	(c) " (a)	.24
(c) E–D (total)	.323	.095	29.41	(b) " (c)	3.30
(d) E–E (different material)	.320	.109	34.06	(b) " (d)	3.16
(e) E–D (same material)	.327	.080	24.46	(b) " (e)	3.13

* Pearson's coefficient of variation.

An examination of the standard deviations of these different groups of correlations and the resultant measures of relative variability sheds some interesting light, in a preliminary fashion, on one of our major problems. The standard deviations and coefficients of variation of the E–E group and the various E–D groups are of about the same order of magnitude, but despite the larger mean of the D–D group, its standard deviation is hardly more than

---

[11] All average correlations presented in this paper were obtained by first transmuting the Pearsonian correlations into Fisher's 'z' (9, 177) and then obtaining the average. Unless otherwise stated, the averages presented have not been retranslated into Pearsonian correlations.

60% of the corresponding measure of the E–E group.[12] This would indicate that the intercorrelations of the difficult variables are far more homogeneous in magnitude than the corresponding intercorrelations of the easy variables. It would suggest that a simpler factor pattern could be made to fit the former set of correlations than the latter. At any rate, independently of any factor theory or analysis, we could conclude from the higher intercorrelations among the difficult variables and their lower variability that, on the average, pairs of variables show a greater relationship to each other when they are difficult variables than when they are easy ones and that the relationships existing between the pairs of the difficult variables are subject to less variability than are the corresponding relationships among the easy variables.

b) Relationships among correlations involving variables of various levels when the effect of material is considered.

The previous discussion was based on an analysis of the effect of the level of difficulty on the correlations that were reported without particular reference to the different types of material.[13] With respect to homogeneity of both material and level we may define four different types of intercorrelations. They are the intercorrelations among the easy numerical variables (nn), easy spatials (ss), difficult numericals (NN), and difficult spatials (SS). For purposes of studying more carefully the effects of introducing various degrees of heterogeneity in the possible combinations of correlated variables, Table V was constructed. Table VI presents the necessary data for evaluating the differences between some of the means that are in the previous table.

The first column represents the intercorrelations among variables alike with respect to both material and level. The second column shows the effect of intercorrelating these variables with others that differ from them only with respect to material. In the third column, the difference is that of level, and in the fourth, it is a difference both of level and material.[14]

---

[12] Ordinarily a direct comparison of standard deviations with one another would be inadvisable. In this case, however, we can be certain that although we know little about the units of the distribution of the two sets of correlations, we know that they are of such a nature that differences in absolute value indicate a true order of differences although the absolute differences may not be measured with complete accuracy.

[13] For purposes of clarity, the following symbolism will be adopted and used whenever possible throughout the paper to represent the various types of correlations and mean tetrads which may be presented: Easy numerical and easy spatial variables will be symbolized by *n* and *s*, respectively, while difficult numerical and spatial variables will be symbolized by *N* and *S*. Correlations and mean tetrads will consequently be symbolized by various combinations of these four letters. For example, a correlation between two easy numerical variables will be indicated by r_{nn} or, simply, by nn; of an easy numerical with a difficult spatial variable by r_{ns} or, simply, by nS; etc.

TABLE V

AVERAGE CORRELATION (z's) AMONG VARIABLES OF DIFFERENT DEGREES OF HOMOGENEITY WITH REFERENCE TO LEVEL AND MATERIAL*

Type of Variables	*Same L*** " *M*	*Same L* *Different M*	*Different L* *Same M*	*Different L* " *M*
Easy Numerical	.259 (6)	.335 (10)	.318 (12)	.265 (10)
Easy Spatial	.440 (3)	.335 (10)	.345 (6)	.376 (10)
Difficult Numerical	.409 (6)	.371 (10)	.318 (12)	.376 (10)
Difficult Spatial	.443 (3)	.371 (10)	.345 (6)	.265 (10)

** L and M refer to level and material, respectively, in the table.

* The numbers in the parentheses represent the number of correlations that were averaged for each group.

If we first examine the average intercorrelations among the homogeneous[15] groups of variables, we find that with the exceptions of the nn group they are virtually of the same magnitude. The latter group (nn) is far below the others, having an average of little over 60% of NN, the next lowest in magnitude. As far as interrelations are concerned, the other groups are far more homogeneous. Since the size of the correlations is here taken as a measure of functional homogeneity, there is no evidence on the basis of the first column for any difference with respect to intra-group homogeneity when the two levels of the spatial test are compared, whereas the difference in this respect between the two sets of numerical tests is quite marked. This may be due to the fact (already discussed in section 1 of the results) that while the two easy levels are both easier than the difficult ones, they are not of equivalent difficulty. The results in the previous section also tend to show that the difficult level is not homo-

---

[14] Correlations involving any of the numerical analogies variables (1, I) with any of the spatial analogies (5, V), and numerical generalizations (2, II) with any part of the spatial generalizations (6, VI) were not used in these calculations since Smith (27) had previously shown that this type of structural similarity was a factor making for higher correlations and our interests lie in studying the effects of material. The 'construction' tests (4, IV, 7, VII) were not eliminated as the designation was almost purely analogical. There was no clear-cut evidence indicating that dividing a figure into parts and continuing a series of numbers were functionally the same act. Since we could not get enough combinations of correlations between variables of like structure and different material with tests 1, I, 2, II, 5, V, 6, VI to justify treating them separately, these combinations were eliminated so that we could study the influences of material and level only. The elimination of the combinations in question did not markedly affect the averages. If they had been used, the greatest change would have been the raising of the average NS correlations .012.

[15] We do not imply functional homogeneity, *i.e.*, that the intercorrelations among these variables are necessarily high.

TABLE VI

EVALUATION OF THE DIFFERENCES BETWEEN SOME OF THE AVERAGES IN TABLE V

*Groups of Correlations whose Differences are Evaluated**	M_1**	M_2**	*D*†	D/σ_D††
nn & ss	.259	.440	.181	4.64
nn " NN	.259	.409	.150	4.69
nn " ns	.259	.335	.076	2.05
nn " nN	.259	.318	.059	1.84
ss " ns	.440	.335	.105	2.37
ss " sS	.440	.345	.095	2.16
ss " NN	.440	.376	.064	1.52
NN " NS	.409	.371	.038	1.46
NN " Nn	.409	.318	.091	2.94
NN " Ns	.409	.376	.033	1.03
SS " SN	.443	.376	.072	2.77
SS " Ss	.443	.345	.098	2.80
SS " Sn	.443	.265	.179	7.16
Sn " Nn	.376	.318	.058	1.61
SN " Nn	.371	.318	.053	2.30

* See discussion of the symbols used in footnote 13.

** M_1 is the mean of the first set of correlations in each pair. M_2 is the mean of the second.

† Absolute differences.

†† Some of these ratios would have been higher had correlation between correlations been considered. With a ratio of 2.00 or over, the chances are 98 in 100 that the differences is greater than zero. For the lower ratios, 1.03, 1.46, 1.52, 1.61, and 1.84, the chances are, respectively, 84, 93, 94, and 96 in 100 that the differences are greater than zero.

geneous with respect to difficulty. It is possible that the easy spatials are not only more difficult than the easy numericals, but that their items are more discriminating. The values of the coefficients of variation are consistent with this possibility, being, or the spatial set, 20.4% and for the numerical set 12.9%. We might expect that an extraordinarily easy test would make our subjects more homogeneous, discriminating less accurately and in all probability lowering correlations.

In reading Table V across and comparing one horizontal line with another, some interesting results can be observed. The easy numerical variables have higher correlations with easy spatial and with difficult numerical variables than they have with each other. The correlations with the easy spatials are higher than with the difficult numericals, but the difference (.017) is too small to be significant. It should be noted that the first finding is directly contradictory to the general conclusions to be derived from the previous analysis of Smith based on the whole variable (easy and difficult items). He found on the basis of using the whole variable that the intercorrelations among numerical groups were slightly but not significantly higher than were the correlations among the cross groups consisting of numerical and spatial variables. Actually, between the two lower levels (from the point of view of difficulty) of the numerical and spatial variables, there is a stronger bond than among certain numerical variables homogeneous both from the point of view of level and, of course, material. We find in this case a functional homogeneity which transcends common-sense reference points. However, the same conditions do not hold when the easy numerical variables are compared with the difficult spatial ones. Here we find that the correlations with the difficult spatial are but .006 above the intercorrelations of the easy numerical. One should, however, consider the factors that in this situation make for positive correlations before trying to fit the latter value into a system. In view of Smith's (27) and Schneck's (26) work, it is likely that material plays a role (though one not independent of level). Thorndike (33) reports that correlations between levels are affected by the relative 'distance' of the levels from each other, being lower, the further the levels are apart.[16] Lastly, we should not neglect the fact that the relative strength of a variable determined from its correlations with like variables may be indicative of a strong general factor which influences the magnitude of correlations in other groupings. If we were to arrange the four groups of variables into an hierarchy according to level, we would, on the basis of previous analyses, place them in the following order of difficulty, from easiest to most difficult: easy numerical, easy spatial, difficult numerical and difficult spatial variables. When arranged in this order, the probability would be that the difference between the easy numerical level and the difficult spatial level is quite remote. If we reorganize the correlations of Table V as shown in Table VII, very interesting results are obtained.

The first line of Table VII represents the average correlations between variables of adjacent levels. The second line presents similar correlations between variables of alternate levels and the third line, the correlations between the variables of the two extreme levels. In the first line, differences in level offer the least barrier so that the influence of the general tendency of a variable determined from a given level may be most effective. The lowest of these cross correlations involves the easy numerical values which we know, from considering their own intercorrelations, to contribute a relatively

---

[16] See the discussion of this point in the discussion of the literature above.

TABLE VII

ANALYSIS OF THE EFFECT OF LEVEL

*Ascending Scale of Difficulty**	*Easy Numerical*	*Easy Spatial*	*Difficult Numerical*	*Difficult Spatial*	*Average***
Average of z's between variables of adjacent levels	.335	.376	.371		.361
Average of z's between variables 2 steps apart		.318	.345		.327
Average of z's between variables 3 steps apart		.265			.265

* These steps are not to be thought of as being separated by equal degrees of difficulty. They merely represent the order of difficulty obtained through an analysis of the success of the group of subjects.

** In computing the averages, the number of correlations involved in each group had to be considered. The only average in the table affected by this is the second one. The number of cases involved in each group are to be found in Table V.

small general tendency to be operative in correlations with other variables. The other two sets are higher and are about equal to each other in magnitude. In the second line, the correlations are lower than in the first, but not very much lower. The separation of levels is wider but it is counterbalanced by the similarity of the material The third line correlations average the lowest. There is no similarity in material and we have reason to believe that the separation in level is the widest possible under the conditions of our experiment.

For the remaining groups, i.e., those except the easy numericals, that are homogeneous from the point of view of material and of level, the various groups of cross correlations are all lower than the intercorrelations with which they are to be compared. In the case of the difficult numerical variables, however, the relationship with the spatial variables of either level is particularly strong. The differences between the averages are not reliable whereas the chances are 99 out of 100 that the corresponding difference between the difficult numericals and the correlations of the difficult numericals with the easy numericals is a true one. The fact that sN and SN correlations are higher than the nN correlations would indicate greater functional homogeneity between parts of the numerical tests and the spatial tests than between one part and another part of the numerical tests.

Similarly, the difficult and easy spatials appear to be related to the difficult numerical variables at least to the same extent that they are related to each other. This is not the case when comparison is made with the easy numerical variables whose correlations average .083 below the cross correlations of easy and difficult spatials.

In evaluating the differences between some of the measures discussed in this section, it was felt that the use of Fisher's 't' method (9, 99 ff) was unnecessary as our chief interest here lay in hierarchies of relationships rather than in determining the exact probability in every case for the extent of the reliability of the difference. Moreover, with some of the groups having as few as three correlations, the correction necessitated by the Fisher technique would be too great and might tend to distort some very real trends.

It will be seen that the intercorrelations among the difficult spatials differ, with greater reliability than the other groups of intercorrelations, from the various cross groups with which they were compared. The reliability of these measures is due not so much to the high average to be found within this group as to its relatively small variability. There is also shown a high probability that the intercorrelations among the easy spatial variables and the difficult spatial variables taken as separate groups differ from the cross correlations among easy and difficult spatials, the two former groups being the higher ones. This is the most unequivocal evidence that may be presented here regarding the influence of level of difficulty as a factor in making for functional homogeneity. The average intercorrelations of each of the two spatial groups are virtually identical, differing only by .003. Yet they differ from the cross correlational averages by .095 for the easy group and .098 for the difficult ones, a difference which when considering the factors which have tended to depress the absolute magnitude of these correlations is by no means a negligible one.[17]

With reference to the two levels of the numerical variables, the cross correlations do not fall below both levels largely because of the wide gap between their average correlations. It is interesting to note, however, that the cross correlation average is .091 below the average of the difficult numerical correlations (28.6% of the cross correlation average). This number corresponds to the differences obtained with reference to the spatial groups, and the intercorrelational average of the difficult numerical variables is of the order of the intercorrelations among the spatials. The nN average, while lifted by the relatively high value of the difficult numerical variable, is not reliably above that of the intercorrelations among the easy numerical variables. The difference is .059 and the criterion ratio 1.84.

---

[17] The differences were 27.5%, and 28.4%, respectively, of the cross correlational average.

Table V also helps to explain the factors involved in forcing the E–D correlations down to the level of the E–E ones. The level is really above that of the intercorrelations of the easy numerical tests and below the easy spatials. The average value of .323 tends more towards the numerical than the spatial correlations. The E–D correlations are depressed by the influence of the correlations among the easy numerical and difficult spatial variables. If they are eliminated the average correlation of E–D rises to .347 which is slightly higher than the average of .329 for the intercorrelations among all the easy variables, but not higher than .357 the figure obtained when the correlations are eliminated. When the material factors are held constant, our comparison becomes nothing more than the comparison between levels which we have already discussed.

The general conclusions to be derived from this section are: (1) If single test variables are subdivided into two tests consisting of the relatively easy and the relatively difficult material, respectively, the average correlations among easy and difficult variables will not be the same as for the intracorrelations within the easy and difficult groups separately. (2) The cross correlations between easy and difficult tests of the same kind of material will tend to be lower than the intercorrelations for the easy and difficult groups separately, provided that within their own group the easy and difficult variable are of about the same strength, i.e., have reached the same level of intercorrelation. (3) When the intercorrelations of the easy and difficult variables, taken separately, are not equal, we may infer that the 'strengths' of the variables concerned are not equal. In this case, while the cross correlations will be below the intercorrelations of the stronger group, it may rise above the weaker group, provided that the strength of the higher group is enough to counterbalance the difference in level. (4) When comparing tests containing different material, it was found that the division into easy and difficult variables gave evidence for conclusions contrary to those that could be drawn when the variable was intact. Between the easy parts of the numerical variables, a stronger bond with the easy spatials was found than was to be found among the easy numerical variables themselves. When the whole variable was used, the numerical correlations were slightly higher than the numerical-spatial correlations. (5) The relationship between the difficult numerical variables and both types of spatial variables was at least as great as the relationship among the two types of spatial variables themselves. With the whole variables, the intercorrelations of the spatial variables tended to be higher than the numerical-spatial correlations. (6) The relationship between difficult numerical variables with both types of spatials is also greater, but not with absolute statistical reliability than are their relationships to the easy numerical variables.

All these latter facts indicate that when the whole variable was used as a unit, it contained within itself heterogeneous factors. Relationships obtained between variables were compromises between the various tendencies that these heterogeneous factors indicated and depended largely upon which were predominant. Conclusions referring to material cannot be made without reference to level.

c) Correlations between easy and difficult parts of the same variable.

The average correlation between easy and difficult parts of the same variable was .546. This value is greater than .399, the value of the average intercorrelations among the difficult variables. The difference of .147 is statistically reliable, the criterion ratio being 3.42. The other broad groups, E–E and E–D, differ more widely in magnitude than does D–D from .546.

However, the easy-difficult correlations of the same variable have a very wide range. Two of these values, those for the arithmetical series, (variables 4, IV) and for the spatial analogies (5, V) are rather low, being equal to .352 and .458, respectively. It is interesting to see whether in terms of correlations the variables 4, IV, 5, and V have stronger relations with other variables than they have with the easy or difficult half, as the case may be, of the variable from which they have been abstracted. We find that this is so in 4 out of the 12 correlations involving variable 5, in 3 out of the 12 for variable V, in no case for variable 4, and in 11 out of the possible 12 cases in variable IV. This overlapping procedure is relatively crude; it is important to find whether the averages of the correlations differ significantly from the correlations with their easy or difficult counterpart. The number of cases being small, (N = 12), but not too small to influence unduly the measurement of the reliability of the differences, and our desire being to determine differences with fair accuracy rather than to find an hierarchy, Fisher's 't' method was used (9).

With variable 5, the chances were 79 in 100 that the difference was greater than zero, the average of the other correlations involving variable 5 being .397. The average for the correlations of variables 5 with the easy variables was slightly higher, being .405, but the decreased number of cases resulted in changing the chances to 4 in 100 that the difference was greater than zero. This adds to our general feeling that with a small number of cases there is a limit beyond which the 't' method may obscure the general tendencies.

With variable V, there were 85 chances out of 100 that the difference was a true one (.389 versus .458), but when the average correlation of V with the difficult variables only (.441) was used, the chances were 2 in 100 that the difference was greater than zero. In this case, the reduction appeared to be more reasonable than in the case of variable 5 as .441 differs from .458 only by .017.

Variable 4 has the lowest correlations with the remaining 12 variables than any of the other variables in this study. The determination of its variance due to its correlations with other variables (to be discussed in a later section) further emphasizes this fact, the value of 18.1% being the lowest obtained. We therefore did not attempt to evaluate the differences of its correlations with respect to the correlation between 4 and IV. Using variable IV, however, we find that its average correlation of .421 is greater than the correlation of .352 (of r_{4IV}); by the 't' method, the chances of the difference being a true difference were 82 out of 100. The correlations of variable IV with the other variables were not subdivided into those with the easy and those with the difficult variables as no difference was found between the latter correlations (.423 for the easy and .419 for the difficult).

In general, these results may be interpreted as indicating that for the four variables involved in the two lowest correlations between halves of the original variables, the correlational averages with other variables were, for three of them (5, V, 4) reliably below the 'same' variable correlation and for one of them (IV) reliably above.[18] For selected sets of correlations involving variables 5 and V, there were statistically no reliable differences though in absolute value the averages were still below the value of the correlation with which they were being compared.

---

[18] The next highest correlation between easy and difficult 'like' variables was .528 (r_{6VI}). In view of the above results which indicated that with the two lowest correlations we had already established boundary conditions, the above type of analysis was not carried further. It should be noted, in justification of this policy, that no correlation other than correlations between easy and difficult parts of the same variable were equal in value to .528. No correlation with variables 6 or VI came very close to this third lowest value, the nearest being .434 ($r_{IV\,VI}$).

These results would indicate that there were few instances of stronger bonds between one variable and another than between a variable and its corresponding easy or difficult complementary variable. There are only 18 such occurrences out of a possible 168, i.e., 10.7% of the times.[19] Even this small figure is chiefly the work of one variable, number IV, which supplies 11 of the 18 instances. If this were eliminated, we would have the occurrence of these stronger bonds in only 4.5% of the instances now under consideration. We must conclude that there are very few cases of a variable having a stronger bond with another variable than it has with its 'counterpart.'

However, there does occur, with greater frequency, bonds, between variables, that are stronger than the lower order of bonds that exist between some of the easy and difficult variables of what were originally the same test. There are 39 instances out of the possible 84 in which the lowest correlation (.352) is exceeded and only 9 instances where the next lowest (.458) is used as a reference point.[20] There is no overlap with the third lowest correlation (.528).

The low value of r_{4IV} is in all likelihood due to the low level of difficulty that variable 4 was sampling. It has the highest percentage of correct answers of all the tests that have been used in this study, and again emphasizes the effect of differences in level on correlations.

3. ANALYSIS THROUGH THE USE OF MEAN TETRADS

The mean tetrad has recently broken into prominence as a measure of relationship. It has been used by Kelley (21), Tryon (38), and Smith (27) among others. A formula which greatly simplifies its calculation has been derived by Brolyer (4). This measure developed with the multiplication of studies in factor analysis. With these studies came a rapid "discovery" of various 'group' factors. The latter were usually manifested by a disturbance of the tetrad equation so that its value had to be regarded as being greater than zero even with the most liberal use of the probable error. The group factor, in its simplest form, may be thought of as an additional bond between two variables which makes for higher correlation between them. When the variables in a study rose to any considerable number, the treatment by simple tetrads alone became almost unmanageable. With the increase of the number of variables would come a more than proportional increase of possible bonds so that detecting them in any strict way became well nigh impossible. This becomes clearer if we examine an illustration. Kelley (21) and Tryon (38) have shown that of tetrads assume certain patterns, these patterns may be indicative of certain bonds. Thus, if $t_{1234} = t_{1243} > 0$, and $t_{1342} = 0$, we may conceive of this pattern as having arisen through the presence of a general factor plus the presence of a group factor between either variables 1 and 2 or between 3 and 4 (or even one between each of the two pairs). In order to find where the bond is, set-ups such as the following are then analyzed: t_{12ab}; t_{34ab}. But in doing so, we cannot solve our problem if there are bonds between 1 and 2 *and* between 3 and 4. This method also results in the introduction of possible additional bonds between the new variables and the original ones. As a matter of fact, the two latter conditions have often become insuperable obstacles to this type of analysis. If the first two variables in the tetrad are kept constant, however, and the next two are varied so as to bring in all possible combinations of the variables under consideration (except, of course, the first two), we get what appears to be a measure of the relative strength between the pairs of variables in question. For the tetrads of the form t_{abij} give all the possible tetrads in which ab may be associated together as a unit. By averaging these tetrads, we obtain the mean tetrad for variables ab, which, in a sense, is thought of as a measure of the relative strength of the pair of variables, ab, as compared to the other possible pairs in a particular set up.

---

[19] Each of the 14 variables has 12 possible correlations with other variables aside from with its easy or difficult 'counterpart'. Each correlation must be considered twice, as a correlation is taken between two variables and, therefore, in this analysis, every one of them must be compared not only with the r_{aA}, but with the r_{bB} correlation.

[20] In this case, N is equal to 84 as the seven E–D correlations with like variables are excluded and no correlation may be counted more than once.

We should, however, bear the following facts in mind. If we expand a typical tetrad equation, we can see that the value of a given mean tetrad depends on more than just the bond between the two variables in question. Thus, $t_{abij} = r_{ab}\, r_{ij} - r_{ai}\, r_{bj}$. The positive part of the mean tetrad, besides being a function of r_{ab}, the correlation between the pair under consideration, is also a function of the remaining correlations. Should a and b have strong bonds with other variables, the fact that they are excluded from the ij combinations (since a, b, i, and j must be four different variables) tends to lower the positive part. If their correlations with the remaining variables are lower than the average, the opposite effect will be found. When we consider the negative part, strong bonds with other variables raise its absolute value, thus resulting in a depression of the total value of the tetrad difference. A low set of intercorrelations with the remaining variables will decrease the negative side and increase the total value of the tetrad difference.

Thus, it would seem to be a dangerous procedure to consider the mean tetrad an absolute measure of relative strength. This danger is lessened if variables tend to have special bonds with only a few other variables. When this is the case, we can usually detect these bonds by the mean tetrad technique. But if a variable tends to have a general 'strength' or general 'weakness,' the mean tetrad may be misleading. The following analyses, utilizing mean tetrads, will therefore be carefully evaluated from this point of view in order to make certain whether or not the results they indicate are valid. Attempts will be made to describe specific conditions which may obscure real relations or indicate false ones, should there be any.

When we consider the results of our analysis of the correlations, we find that the possible set ups within which we may determine our mean tetrads are very numerous. Offhand, it would seem advisable to use the entire correlation matrix. Should we do so, however, we may conceal many opposed tendencies which, in this general treatment, may nullify each other. Moreover, if we were to use all the correlations, it would be difficult to eliminate the influence of correlations between easy and difficult variables derived from the same test. Consequently, mean tetrads were derived from various groups of correlations in order to find whether they shed further light on the relationships already determined. Mean tetrads were found, using (a) only the easy variables, (b) only the difficult variables, (c) easy numerical variables with difficult spatial, (d) difficult numerical variables with easy spatial, (e) additional set ups for mean tetrads for difficult variables where ij were easy variables; for easy variables where ij were difficult variables; and for easy difficult combinations, where i was an easy variable and j a difficult one and vice versa. We shall see later that the distinction between the two latter groups is an important one. In general, the tetrads within each set up were divided into four kinds. They were mean tetrads for pairs of numerical variables; for pairs of spatial; for pairs of numerical plus the spatial pairs (i.e., between variables that were alike in material); and for pairs consisting of one spatial and one numerical variable (i.e., pairs unlike in material). This enables us to evaluate the differences between tetrads alike in material and those composed of variables of different material. We were also able to find whether there was a difference between the spatial and numerical variables as shown by the mean tetrad.[21] Since so many mean tetrads had to be calculated, most of the results are presented in terms of the averages of the important groups defined above. A typical table involving mean tetrads derived from the difficult variables alone is presented below. It is to be read like a correlation table, the subscripts of a particular mean tetrad being the number of the row and column, respectively, of its cell.

Table IX presents the data for the evaluation of important averages of mean tetrads taken from the various sources discussed above.

TABLE VIII

A Typical Mean Tetrad Table (t_{ABIJ})

Variable	*I*	*II*	*III*	*IV*	*V*	*VI*	*VII*
I		.0220	–.0167	–.0188	.0205	–.0143	.0073
II			.0353	–.0009	–.0229	.0097	–.0457
III				.0347	–.0041	–.0350	–.0117
IV					–.0218	.0157	–.0062
V						–.0033	.0266
VI							.0298
VII							

Underlined mean tetrads are those between numerical analogies and spatial analogies; and numerical generalizations and spatial generalizations. These have not been used in calculating the averages of groups of mean tetrads.

a) Analysis of tetrads of the form t_{abij} (involving the intercorrelations among the easy variables).

The results from Table IX indicate that the relations obtained from analyzing the mean tetrads tend to distort the general relations that were previously indicated by the correlations. There is no difference that is statistically reliable when using sets of like tetrads to be compared with cross tetrads, although the result obtained when using the spatial mean tetrads verge on a truly reliable difference. The numerical bonds, although they are weak, are stronger than the cross group bonds. In general, we would conclude, on the face of the results from the mean tetrads of this set up, that material is not very influential in acting to create bonds, although in all cases mate-

---

[21] As in the case of the correlations, tetrads involving combinations of variable 1, I with 5 or V and 2 and II with 6 and VI were eliminated. This resulted in decreasing most of the cross tetrad groups by two combinations only. The influence due to the similar structure of the analogies tests and the generalizations tests would probably be greater here than when just correlations were considered, as their influence would be multiplied by the number of tetrads involved in making up the mean tetrad between them.

TABLE IX

EVALUATION OF THE DIFFERENCES BETWEEN THE AVERAGES OF VARIOUS GROUPS OF MEAN TETRADS*

Source	*Types of Mean Tetrads*	M_1**	M_2**	$(M_1 - M_2)$ D	D/σ_D
Easy num. &	(nn & ss) vs. ns	.0011	–.0058	.0069	1.17
easy spatial	nn “ ns	–.0006	–.0058	.0052	.70
	ss “ ns	.0044	–.0058	.0102	1.50
	nn “ ss	–.0006	.0044	–.0050	– .40
Diff. num. &	(NN & SS) vs. NS	.0121	–.0139	.0260	2.95
diff. spatial	NN “ NS	.0093	–.0139	.0232	2.17
	SS “ NS	.0177	–.0139	.0316	3.10
Diff. num. &	(NN & ss) “ Ns	.0100	–.0089	.0189	1.64
easy spatial	NN “ Ns	.0060	–.0089	.0149	1.12
	ss “ Ns	.0181	–.0089	.0270	2.48
Easy num. &	(nn & SS) “ nS	.0172	–.0086	.0258	3.44
diff. spatial	nn “ nS	.0129	–.0086	.0215	3.12
	SS “ nS	.0258	–.0086	.0344	3.70
Mean tetrads	(NN & SS) “ NS	.0249	.0164	.0085	.91
of the form	NN “ NS	.0145	.0164	–.0019	– .17
t_{ABJI}	SS “ NS	.0457	.0164	.0293	3.57
Mean tetrads	(nn & ss) “ ns	.0257	.0201	.0056	.66
of the form	nn “ ns	.0233	.0201	.0032	.37
t_{abIJ}	ss “ ns	.0307	.0201	.0106	1.09
Mean tetrads	(nN & sS) vs. Unlike‡	.0038	–.0016	.0054	.70
of the form	sS “ “	.0063	–.0016	.0079	.69
t_{aBJI}	nN “ sS	.0023	.0063	–.0040	– .27
Mean tetrads	(nN & sS) vs. Unlike‡	–.0127	–.0262	.0135	1.73
of the form	nN “ “	–.0063	–.0262	.0199	2.46
t_{aBIJ}	sS “ “	–.0237	–.0262	.0025	.21
	(nN & sS) “ nS	–.0127	–.0310	.0183	2.77
	“ “ sN	–.0127	–.0214	.0087	.89
	nN “ nS	–.0063	–.0310	.0247	3.58
	nN “ sN	–.0063	–.0214	.0151	1.47

* Each mean tetrad on which the averages have been found is itself an average of 20 individual tetrad differences. With the exception of groups derived from t_{aBJI} and t_{aBIJ}, the number of mean tetrads of numerical material was 6, of spatial 3, and of numerical spatial 10. For the above mentioned mean tetrads, the corresponding number of cases is double, being 12, 6, and 20, respectively.

** This represents the order of the means of the groups as given in the preceding column.

‡ This represents the average of the total unlike (*i.e.*, unlike in material) mean tetrads.

rial bonds would seem to be stronger than bonds between variables not enjoying similarity of material. When we return to our fundamental table (Table V), we find that some of these results are contradicted by it and are due to the misleading character of the mean tetrad. The average of the correlations between easy spatial and easy numerical variables is actually .076 above the intercorrelations between the easy numerical variables, the criterion ratio of this difference being 2.05 (i.e., there are 98 chances out of 100 that the difference is greater than zero). The spatial correlations are considerably above the cross correlations (.105 above). The critical ratio of 2.37 indicates that there are 99 chances out of 100 that the difference is greater than zero. Most interesting of all is to find that the criterion ratio between easy numerical tetrads and easy spatial tetrads is only .40 the difference between the means of these mean tetrads being only .0050) indicating no real difference; while actually the correlations involving these groups differ on the average by .181 with a critical ratio of 4.64.

We can see, from the analysis that follows, how some of these contradictory results have arisen. For any mean tetrad between two numerical variables we must average tetrads of the form $t_{n_1n_2ij}$. For 'spatial' tetrads, we will have a similar expression and for mean tetrads between spatial and numerical variables, we average tetrads of the form t_{nsij}. It is worth our while to make a careful analysis of factors other than just the n_1n_2, s_1s_2, and the ns bonds which enter into the particular relationship indicated by each of the three kinds of tetrads which supposedly measure the relative strength of these bonds. Maintaining our usual notation, such an analysis is given in Table X where we have presented the correlations other than the nn or the ns which affect the value of the mean tetrad for a particular nn or ns bond.[22]

If we compare the positive sides of the two kinds of tetrads, we find that each of them has an equal number of correlations between numerical and spatial variables. The numerical tetrad has two correlations between numerical variables as against six easy numerical correlations for the spatial tetrad. With regard to correlations involving only spatial variables, the numbers are reversed. Since numerical correlations are considerably lower than spatial, in this set up, the aforementioned facts mean that factors, other than the bonds in question, tend to make a bond between two easy numerical variables appear stronger than a bond between an easy spatial numerical variable. A similar situation which makes the negative part of the numerical-spatial tetrad larger (and consequently the entire tetrad difference smaller) and the corresponding part of the numerical-numerical tetrad smaller exists as is shown by the table. The two types are equated for the number of r_{ns} r_{ns} combinations.

---

[22] The comparison between the factors entering into the n_1n_2 and the ns type of bonds only are presented, as the chief purpose of this analysis is illustrative. Similar relationships may be shown for s_1s_2 bonds, but it was thought unnecessary to analyze them specifically.

TABLE X

ANALYSIS OF THE CONTRIBUTION TO THE VALUE OF THE MEAN TETRAD OF THE FORM t_{abij} OF CORRELATIONS OTHER THAN THE ACTUAL CORRELATION BETWEEN a AND b

	Postive Part of Mean Tetrad		*Negative Part of Mean Tetrad*	
	t_{n1n2ij}	t_{nsij}	t_{n1n2ij}	t_{nsij}
Types of r's other than r_{ab} that are involved	$2r_{nn}$	$6r_{nn}$	$2r_{nn}r_{nn}$	$6r_{nn}r_{ns}$
	$12r_{ns}$	$12r_{ns}$	$12r_{nn}r_{ns}$	$6r_{nn}r_{ss}$
	$6r_{ss}$	$2r_{ss}$	$6r_{ns}r_{ns}$	$6r_{ns}r_{ns}$
				$2r_{ns}r_{ss}$

But the numerical-spatial mean tetrad has no r_{nn} r_{nn} combination, whereas the other has two of them. The numerical-spatial has only 6 r_{nn} r_{ns} correlations, while the numerical-numerical has 12. Corresponding to the six remaining r_{nn} r_{ns} combinations for the numerical-numerical mean tetrad, we have 6 r_{nn} r_{ss} combinations which will be larger magnitude than the corresponding r_{nn} r_{ns} measures. Of course, the 2 r_{nn} r_{ss} combinations of the numerical-spatial tetrad will be larger than the 2 r_{nn} r_{nn} combinations of the numerical-numerical. The total effect will be that those factors independent of the particular bond studied will go toward raising the absolute magnitude of the negative part of the cross group mean tetrad above that of the corresponding part of the numerical-numerical mean tetrad. The bond shown by the tetrad is thus partly a function of spurious factors. In this situation they tend to reverse the true relation regarding the strength of bonds between easy numerical variables as compared with bonds between easy spatial and easy numerical variables. A similar analysis could be made to show why the really marked difference between the strength of easy numerical bonds as compared to easy spatial bonds is completely obscured by the mean tetrad technique.

b) Analysis of mean tetrads of the form t_{ABIJ} (involving the intercorrelations among the difficult variables).

From the figures presented in Table IX, the chances seem rather high that mean tetrads consisting of variables of the same material are reliably larger than cross group mean tetrads, with the spatial group again showing the widest difference. These differences are in the same direction as those indicated by the relevant correlations to be found in Table V. One should note that while these differences do not contradict the correlation coefficient, they tend to reveal another weakness of the mean tetrad technique, as we cannot interpret from the mean tetrad the true magnitude of the difference of the relationship. A large probability of a difference greater than zero may be obtained by using mean tetrads simply because of a certain consistency in the trend of the individual tetrads that make them up. When we return to Table V we find that the difference between the intercorrelations of difficult numerical variables and the difficult cross group variables is only .038. One of the defects of the mean tetrad is that it multiplies the same situation so many times that insignificant differences in relationships may appear significant, or, nearly significant, statistically. If N's and S's are substituted for the small ones in Table X, we may again study the influence of extraneous factors that raise the difficult numerical mean tetrads and lower the numerical spatial ones. This effect is, however, not so marked as it was in the first case because of the greater uniformity of the correlations under consideration.

c) Mean tetrads with the easy numerical and difficult spatial variables.

The evaluation of the reliability of the differences between like and unlike groups is here of the greatest interest. The criterion ratios are all greater than three, including the one involving the numerical-numerical tetrads. We know, however, that there is substantially no difference between the intercorrelations of the easy numerical variables and the cross correlations of the easy numerical and the difficult spatial variables. We must look for other reasons and they are to be found in factors similar to those shown in Table X for the mean tetrads of the t_{abij} form. The easy numericals have more strong combinations associated with them on the positive side than do the cross mean tetrads, while on the negative side this relationship is reversed. The tetrads discussed above are a flagrant example of how special circumstances may make the mean tetrad a most misleading measure.

d) Mean tetrads using difficult numerical and easy spatial variables.

In this set up, the mean tetrads align themselves in the same manner that the correlations do. The difference in average correlation among the difficult numerical variables is only .033 higher than the cross group correlations. This is reflected in the low criterion ratio of 1.12 to be found with reference to the corresponding mean tetrads. When the spatial group is used, the ratio of 2.48 implies 99 chances out of 100 that there is a true difference, but, of course, it sheds no light on the significance of this difference.

e) Miscellaneous combinations of easy and difficult variables.

When we consider mean tetrads consisting of two easy and of two difficult variables, we find that the following tetrads may be devised: t_{abIJ}, t_{ABij}. When we use variables of different level as the ones between which the bond is to be found, we find that tetrads of the form t_{aBiJ} cannot be indiscriminately averaged with tetrads of the form t_{aBJi} as the two imply very different relationships. While the positive part of the expanded equations are equal to each other in the two cases, we find that the negative part is, in the first case, $r_{ai}\ r_{BJ}$ while in the second, it is $r_{aJ}\ r_{Bi}$. These imply two different relationships. The negative side of the first kind of mean tetrad will be obtained by averaging the sum of the products of easy-easy correlations with difficult-difficult ones. The negative side of the other is obtained by averaging sums of products of cross correlations. We would expect that the average of the mean tetrads of the first set would be a negative value reliably below zero, while the average of the other would be about zero. This is due to the fact that the product of E–E and D–D correlations will in general be higher than products of pairs of E–D correlations. (See Table IV. This table shows on the average that E–E correlations are about equal to E–D correlations, while both kinds average less than the average D–D correlation.) In the second case, we will be getting E–D products against E–D products and hence the conclusion stated above.

1) *Analysis of mean tetrads of the form* t_{ABij}.

In this set up, only the spatial mean tetrads differ reliably from the cross mean tetrads. The numerical group is actually slightly lower than the unlike tetrads. The corresponding numerical correlations, however, average .409 while the cross group average is .371. While we have an inversion of the correlational relationship when we use the mean tetrad, we may see that this reversal is not so great as might seem since the two sets of correlations are close to each other in magnitude.

2) *Analysis of mean tetrads of the form* t_{abIJ}.

No grouping here shows anything like a reliable difference. Again the relationships involving easy numerical mean tetrads are distorted, this time by the magnitude of the associated IJ correlations. The numerical correlational average is .076 correlation points below the cross group average. In terms of mean tetrads they actually are slightly higher than the cross group tetrads. In addition, the rather large difference of .105 between the spatial correlations and the numerical-spatial group is not truly represented by the mean tetrads obtained for these groups.

3) *Analysis of mean tetrads of the form* t_{aBJi}.

As was predicted above, the average of these tetrads verges on zero. Their actual average is .0010, the standard error of the mean being .0039. The critical ratio between the mean and its standard error is .26, showing conclusively that the true difference is not greater than zero. All the differences obtained between groups of mean tetrads in this set up also have a low probability of being greater than zero. Those presented for the evaluation of the reliability of the difference were groups between which the differences were largest. There was no point in continuing the various subgroup analyses.

4) *Analysis of mean tetrads of the form* t_{aBiJ}.

These mean tetrads present a variety of interesting groups for comparison. The highest reliability index is obtained with regard to the difference between the nN mean tetrads (which are higher) and the nS mean tetrads. This is in conformity with our analysis of the correlations. But there must be other factors that assist in making the difference between the averages of these mean tetrads so large, as in terms of correlations the difference between the corresponding groups is not particularly striking. The difference between the total 'like' and total 'unlike' tetrads conforms to the correlation set up. The numerical mean tetrads are relatively stronger than the total set of unlike mean tetrads. This does not conform to the correlations, as we can find no substantial difference between a value of .318 and .327 (the latter value being the average of all the correlations differing in both material and level).

When the mean tetrads between variables of both different material and different level are subdivided into those between easy numerical variables and difficult spatial (nS) and into the remaining group of easy spatial and difficult numerical (sN) the differences with regard to the total like tetrads change in the expected direction. Using the first group of cross variables, we find that the criterion ratio is raised from 1.73 to 2.77, whereas with the second it is lowered to .89. This is in harmony with the analysis of the correlations, but it does not tell the entire story. The average of the total 'like' correlations in this set up is .327. For the correlation among the easy spatial variables and the difficult numerical variables, it is .376. The mean tetrads reverse the order of difference. For the sS comparison the relationships indicated by the correlations are not distorted by the mean tetrad. Finally, the direction was distorted when nN was used in comparison with the sN combinations. The former mean tetrads were higher despite the fact that the average correlation was lower than the cross group correlation.

The presentation of the t_{aBIJ} mean tetrads indicate clearly that the magnitude of the relationships between variables and the direction of the differences between relationships may be distorted by them. No attempt was made, as in the case of t_{abIJ}, to analyze in detail the particular spurious factors that were involved in each instance. In view of the previous analysis, and the results obtained, it is obvious that they were operating. The manner of operation may be deduced in detail by expanding the mean tetrad equation and studying the implications of its correlational components for the particular relations in question.

In summarizing the discussion of the mean tetrads up to the present point of the presentation, the following statements should be noted. While the use of tetrads, and their derived measures mean tetrads, may have certain implications in Spearman's particular variety of factor theory, we find that when the mean tetrad is used to measure the relative strength of the bond between various pairs of variables it has no particular advantage over the correlation coefficients and in many instances shows at a disadvantage to it. This measure of relationship becomes a function of factors other than the specific bonds studied, and often these factors may be influential in distorting the true order of relationship. It is probable that with wide differences in degrees of relationship, this effect is not so widespread, but even under those circumstances the mean tetrad could hardly tell us more than a careful analysis of the correlation coefficient would.

f) Relationships between variables of the same material and different level.

While the previous analyses were made particularly with reference to relationships among different groups of variables with reference to both material and level, no mean tetrads were presented which were determined from set ups of variables which differed from each other only with respect to level. When we consider the nature and the number of the variables in this study, we shall see that mean tetrads may not be determined from them. Because of the special bonds that exist between easy and difficult variables that were derived from the same test, we know that if we permitted the introduction of such variables into the same tetrad, our results would be hopelessly distorted by this special relationship. If we were interested in finding the relative strength, for instance, of the bond between variables 1 and II by the mean tetrad technique, we would find that using variables I and 2, in a particular determination, would unduly increase the magnitude of the negative part. If we consider set ups with an equal number of easy and difficult variables so selected that no two have been derived from the same original test, the number of tetrads that could be used for evaluating a particular mean tetrad would become very limited. For the spatial variables, there is, of course, no possibility for any tetrad equation under these conditions (since there were originally only three spatial tests). For the numerical pairs, the number of measurements becomes just two for each one. If the implications of the restriction thus imposed are studied carefully they show again how the tetrad and the mean tetrad may be very inadequate measures. If any two easy variables are selected, then the two difficult variables associated with them will have the same value when the average tetrad is calculated. In this special system, bonds between easy pairs automatically become the same as bonds between certain other difficult pairs. This is so because the relationships are of such a nature that each easy pair has a difficult pair correlated with it(and vice versa). Thus, the bond between variables 1 and 2 must become the same as that between III and IV. The bond between 3 and 4 becomes the same as that between I and II, and so on. These facts make it even more difficult to use the tetrad technique legitimately than in the previous cases.

4. ANALYSIS BY THE USE OF TRIADS AND BY AN ITERATIVE FACTOR METHOD

The previous section pointed to the fact that mean tetrad analyses could at best be of about the same value as analyses of the correlation coefficients. It would be desirable, however, to make our analyses in such a manner that the relative strength of a particular variable studied would be measured by comparing its correlations with the remaining variables. Because of the mathematical interdependence of the variables in a correlation table upon each other, it would be advisable to use some form of factor analysis wherein the factor weights might be taken as a more refined measure of the relative strength of a variable under consideration than the raw correlation coefficients themselves. It is fairly certain, at any rate, that the factor type of analysis would be less likely to go contrary to the indications of the correlations than the analyses of the mean tetrad, as the factor weights are more directly dependent on the particular correlations showing the actual relations in question. The remaining correlations of the matrix really become a sort of variable reference point (the weights of a factor for given variables and the remaining correlations, that is, those not involving the given variables would vary inversely, again cutting down on the possibility of internal contradiction), and would not have the direct influence on the measure of relationship that they do have in the case of the mean tetrad.

Out of the various methods developed (those of Spearman (30), Hotelling (19), Kelley (22), and Thurstone (35 and 36)), it was felt that a method first popularized by Spearman (See Holzinger (18, 16ff)) and an iterative method which, bears an algebraic relationship to Hotelling's method[23] would be adequate for our purpose. In Kelley's "Crossroads in the Mind of Man," there are presented several sets of necessary conditions which must be fulfilled in order that one may represent the variance of a group of test variables in terms of particular factor patterns. Proposition V, one of the many referring to the patterns which may be set up for three variables, is of particular interest to us. "Three variables may be thought of as arising from one general factor plus three specific factors when the absolute value of the product of every two of the three intercorrelations is less than the third, and when one or three of the intercorrelations are positive." (21, 41).

This propositon, together with Spearman's work, furnishes us an opportunity to obtain a general measure of the variance of our variables. If we consider the correlation between two variables, we find that it may be expressed in terms of any numbers of factors that we desire. Moreover, with only a knowledge of one correlation, we have no way of ascertaining which of the two variables is 'stronger.' The minimum conditions for making an estimate of the latter sort require three variables. If the conditions of Kelley's fifth proposition can be met, we have a situation wherein the relative strength of a variable with reference to two others may be measured with the minimum assumptions concerning the bonds that relate the three variables to each other. If our correlational values are of such a relative magnitude as to permit this treatment, we may permute our variables so as to measure the 'central' factor for each variable in each set up. An average of such determinations may be taken as a good measure of the relative strength of the variables so treated. This is what is really essential in Spearman's method of measuring variance, the mathematical derivation of which is clearly given in Holzinger (18). Our use of this measure may be met with two objections. It is possible that not every triad will satisfy the conditions of proposition V; also this treatment disregards the influence of group factors.

---

[23] In a recent paper, Alexander (1) has presented a factor analysis of 9 variables, using both Hotelling's method (19) and one of Thurstone's (35). While the values of the factor loadings do not correspond exactly when determined by the two methods, the relative standing of the weights of each variable is not greatly altered when one or the other method is used. The present writer calculated the rank difference correlation between the loadings determined by each method for each of the first two components. The correlation between the loadings of the first component was .85; between the loadings of the second component .80.

A survey of the correlation table (Table III) shows that the conditions of proposition V are fulfilled. All the correlations in this table are positive. The other conditions may be tested by multiplying the two largest correlations in a given triad and seeing whether this product is smaller in value than the remaining correlation. If this is so, then the conditions will, of course, be fulfilled by any other selection within the given triad. In looking for possible combinations which do not fulfill the conditions, we need but find the smallest and largest correlation of each variable with the others. In this way, we will automatically determine three variables to be used as a test.[24] The smallest correlation in the entire table is equal to .1497, between variables 4 and VII. Applying the criterion in the two triads which fulfill the condition just stated (4, 5, VII; 4, V, VII) we find that the criterion is satisfied. .1497 is greater than .0926 and .1189, the results obtained by multiplying the remaining correlations of the respective triads. The conditions are also fulfilled for the next smallest correlation of .1678. It might be objected that the new correlations introduced might be small also. It was calculated, however, that for the product of the 'high' correlations chosen and any other correlation, to exceed the low correlation we would need a value for the new correlation which was introduced that would in most cases be higher than any present in the table.

With regard to the second possible objection, the following should be noted. In this analysis, we are interested in a measure which is an indication of the average influence of a given variable. This influence is expressed not only by mathematical factors of a general nature (that is, conceived as running through all the variables), but by special bonds between different variables. The method of triad analysis enables us to include the influence of these special bonds in a single measure without weighting them unduly.

---

[24] Correlations of the form r_{aA} were, of course, not used in any of these determinations.

As discussed earlier, reliability coefficients were not obtained in this study. In the Hotelling system of analysis, the reliability coefficients assume an important theoretical role as they express the total variance of the factors. Therefore we could not use a method which was truly Hotelling's method of principal components. Since, however, we desired to have set ups within which correlations of particular variables could be compared with the remaining correlations (in addition to the triad method), weights were determined for the variables by using raw correlations and no reliability coefficients and applying an iterative method to the resulting correlational matrix. The weights are therefore determined on the same constants as the triad variances. It is interesting to see whether the two methods will give similar results.

The matrices of correlations used will have zero diagonals and consequently their geometric interpretation cannot be the same as that for the matrices used in Hotelling's method. It is of value to see how the sums of the columns of the matrices would differ in our treatment from the corresponding sums employed in the method of principal components.

When the reliability coefficient is eliminated, the sums of the columns of the correlational matrix are considerably reduced (by $n \times 1$, and also by the fact that uncorrected coefficients will necessarily have to be used). This lowers the weights, but makes them more directly a function of the actual correlations obtained. The elimination of 1.00 from the total correlational sum of each variable does not distort the relative standing that would be obtained if reliabilities equal to one were used. We have reason to believe (See procedure) that the differences in reliability between the variables that could be obtained from the set up we have would not be very great and, therefore, there would not be a great differential effect if corrected correlations were used. In general we may say that the elimination of the reliability coefficient tends to lower the absolute magnitude of the weights obtained, and partially overcomes the tendency which decreases the variability of the distribution of the weights. By eliminating reliabilities, we also make the results of these analyses more comparable to those of the triads which, too, have been obtained without reliability coefficients.[25]

---

[25] This analysis shows the algebraic relationship that our modified method bears to the true Hotelling method. It should not be taken to imply that there exists a geometric relationship between the two.

a) Analysis of combinations using various combinations of seven variables each.

Table XI presents the average variance[26] determined by the triad method for each of the fourteen variables. These variances have been obtained from each variable in three different types of set ups. In one set up, they have been determined from all the variables of the same level of difficulty, in the next, two of the variables are of one level and the other variable of the other, while the third set up is the reverse of the second.

TABLE XI

DETERMINATION OF AVERAGE 'VARIANCE' FOR EACH VARIABLE BY TRIADS*

Var.	*abc*	*aBc*	*aBC*	*Av.*	*Var.*	*ABC*	*ABc*	*Abc*	*Av.*
1.	19.1	22.2	19.0	20.1	I	37.1	34.1	25.5	32.2
2.	27.2	25.5	19.3	24.0	II	39.4	49.3	39.5	42.7
3.	40.4	39.1	30.1	36.5	III	32.3	32.3	26.1	30.2
4.	19.5	19.7	15.0	18.1	IV	42.4	51.1	46.8	46.8
5.	50.7	47.4	36.9	45.0	V	48.6	45.1	35.7	43.1
6.	32.5	31.5	25.9	30.0	VI	34.5	29.5	21.2	28.4
7.	40.3	42.0	35.4	39.2	VII	35.2	30.5	21.9	29.2
Av. num.	26.6	26.6	20.9	24.7		37.8	41.7	34.5	38.0
Av. spat.	41.2	40.3	32.7	38.1		39.4	35.0	26.3	33.6
**Av. num./Av. sp.	64.6	66.0	63.9	64.8		95.9	119.1	131.2	113.1

* The small letters represent easy variables in the triad combinations, while the capital letters represent difficult ones.

** These ratios are expressed in percent.

It is interesting to see how these variances reflect, and in some cases clarify, the findings determined from the original analysis of the correlations. In the easy variables, the variance for the spatials is at all times higher than the variance of the numerical variables. When a difficult variable is introduced into each triad, the spatial maintains its superiority, but the values approach numerical ones more closely. This is due to the fact that the easy spatial correlations with both types of difficult variables are lower than their intercorrelations, while the numerical correlations are raised by being combined with difficult numerical variables and remain about the same when correlated with difficult spatial variables. With two difficult spatial variables, both sets are reduced, but the order of strength remains the same. The reduction is brought about not only by factors mentioned above, but also by the fact that the variances are an inverse function of the correlations between the second and third variables in the triad. When these are both difficult variables, the correlations are higher than when one is a difficult and the other an easy variable. Hence the order of the values obtained. When variables are considered individually, the results would indicate that variable 3, 5, and 7 have the greatest strength when this strength is measured either by averaging all the variances or using only those obtained from the triads consisting entirely of easy variables.

---

[26] The average of all determinations of variance for a particular variable, using all possible pairs of the remaining variables in a particular set-up.

Consistency with the correlation results is also found in connection with the difficult variables. The average variance from the ABC combinations of numerical and spatial variables differ, very little from each other though the spatials are slightly higher. With the introduction of one easy variable, the relationship is reversed due to the low correlations existing between easy numerical and difficult spatial variables. When two easy variables are used, the values of the spatials fall still lower. The depressing influence of the nS correlations are shown in the general averages where variables VI and VII are still very low.

When our modified factor method is used close correspondence is found in comparable groupings. Table XII shows the factor loadings for four set ups. The set ups are those involving all the easy variables; all the difficult; the easy numerical and difficult spatial; the difficult numerical and the easy spatial. In all cases only two components were determined, since they were sufficient to exhaust most of the correlation in the matrix. The results are expressd in terms of the weights of the factors for each variable. In addition, the variance due to the first two factors for each variable and the variance determined by the triad technique are presented.

Again we find that these results are consistent with our original correlation table. For the easy variables, there is a difference between the numerical and spatial variables in the first component. The second component shows the difference to a more marked degree. In general, we shall expect that differences will be shown more decidedly in the second component as the first component tends to be of such a nature as to level out some of the differences. With the difficult variables, the differences between numerical and spatial variables average only .007; for the second component, the average difference is .012. With easy numerical and difficult spatials, the differences again become marked, while with difficult numerical and easy spatials they are again negligible.

When the variances determined by the triads and by the sums of the variances of the first two factors are compared, we find very small

TABLE XII

ANALYSIS OF COMBINATIONS OF EASY AND DIFFICULT VARIABLES BY ITERATIVE FACTOR AND TRIAD METHODS*

Variable	K_1	K_2	$V_{(T)}$	$V_{(F)}$
1	.430	.109	19.1	19.6
2	.481	.128	27.2	24.5
3	.577	.283	40.4	41.3
4	.433	.109	19.5	19.9
5	.613	.311	50.7	47.3
6	.525	.199	32.5	31.5
7	.578	.290	40.3	41.8
Av. spat.	.572	.267	41.2	40.2
Av. num.	.480	.153	26.6	26.3
Av. spat. – Av. num.	.092	.114	14.6	13.9
Av. num. / Av. spat.**	....	....	64.6	65.4
I	.565	.220	37.1	36.8
II	.579	.204	39.4	37.7
III	.540	.180	32.3	32.4
IV	.589	.232	42.4	40.1
V	.616	.261	48.6	44.8
VI	.553	.199	34.5	34.5
VII	.555	.204	35.2	35.0
Av. spat.	.575	.221	39.4	38.1
Av. num.	.568	.209	37.8	36.8
Av. spat. – Av. num.	.007	.012	1.6	1.3
Av. num. / Av. spat.**	....	....	95.9	96.6
1	.421	.083	18.8	18.3
2	.468	.089	25.3	22.7
3	.497	.140	29.4	26.7
4	.418	.071	18.3	18.0
V	.574	.306	43.2	42.3
VI	.526	.257	26.8	33.3
VII	.515	.282	29.9	34.5
Av. spat.	.538	.282	33.3	36.7
Av. num.	.451	.096	23.0	21.4
Av. spat. – Av. num.	.087	.186	10.3	15.3
Av. num. / Av. spat.**	....	....	69.1	58.3
I	.512	.161	27.9	28.8
II	.588	.214	39.9	39.2
III	.549	.210	33.5	34.6
IV	.613	.262	46.8	44.4
5	.618	.279	49.2	45.8
6	.508	.166	27.0	28.6
7	.585	.238	40.3	39.9
Av. spat.	.570	.228	38.8	38.1
Av. num.	.566	.212	37.0	36.8
Av. spat. – Av. num.	.004	.016	1.8	1.3
Av. num. / Av. spat.**	....	....	95.4	96.6

---

* K_1 and K_2 are the weights of the first two components, respectively. $V_{(T)}$ and $V_{(F)}$ are the variances determined by the triad method and from the factor weights, respectively.

** Only ratios between variances have been obtained. Those variances obtained through use of the factor weights adequately represent the tendencies of those weights.

differences. The average absolute difference for the E–E set up is 1.5% in variance; for the D–D, 1.2%; for the easy numerical and difficult spatial, 2.7%; and for the difficult numerical and easy spatial, 1.5%.[27] The variances determined from the Hotelling analysis are slightly lower than those determined from the triads. With a third component, the two sets of variances would very likely be almost identical with one another.

b) Analysis of functional relationships.

Previous results have indicated the possibility of stronger relationships between some variables that differ in material than those that exist between some variables that are similar in material. The analysis of the correlations has indicated in a rough way that the general 'strength' of a variable, by which we mean the general level of its correlations with other variables of a like kind, may be an important factor in making for strong cross relationships.

Accordingly, the weights and variances of each of the fourteen variables were ascertained by deriving them from groups which were the most uniform ones for each variable from the point of view of both level and material. With this as a measure of relative strength under the most favorable circumstances, it would be possible to find whether the variables maintained the same order when more heterogeneous material was introduced.

Table XIII presents the first factor weights, as determined by the modified factor method, and the variances, as determined by triads. The order of relationship obtained by using the two methods are

TABLE XIII

DETERMINATION OF THE 'RELATIVE STRENGTH' OF EACH VARIABLE FROM COMBINATIONS OF VARIABLES HOMOGENEOUS WITH RESPECT TO BOTH LEVEL AND MATERIAL

Variable	K_1	$V_{(T)}$	*Variable*	K_1	$V_{(T)}$
1	.411	19.2	I	.513	31.8
2	.467	33.4	II	.559	45.4
3	.465	32.5	III	.537	37.7
4	.416	19.9	IV	.545	40.6
5	.538	46.6	V	.535	45.4
6	.498	30.6	VI	.512	35.2
7	.541	48.6	VII	.536	45.5

---

[27] These values differ slightly from those that will be found in comparing the differences in the table. This is due to the fact that in the table we have differences between average variance, whereas above we have reported the average of the absolute differences between variance determinations for each variable.

TABLE XIV

THE EFFECT OF THE RELATIVE 'STRENGTH' OF NUMERICAL VARIABLES ON BONDS WITH SPATIAL VARIABLES†

Combinations of variables	*Averages*	K_1	K_2	$V_{(r)}$
1, 4, 5, 7	Spat.	.511	.342	48.4
	Num.	.392	.196	17.0
	D*	.119	.146	31.4
	R**			35.1
2, 3, 5, 7	Spat.	.572	.314	48.0
	Num.	.505	.230	32.7
	D	.067	.084	15.3
	R			68.1
1, 4, V, VII	Spat.	.507	.314	41.5
	Num.	.371	.123	14.0
	D	.136	.191	27.5
	R			33.7
2, 3, V, VII	Spat.	518	.298	40.9
	Num.	.452	.194	24.4
	D	.066	.104	16.5
	R			59.7
I, III, 5, 7	Spat.	.562	.304	47.4
	Num.	.504	.231	30.5
	D	.058	.073	16.9
	R			64.3
II, IV, 5, 7	Spat.	.576	.302	46.2
	Num.	.547	.259	38.4
	D	.029	.043	7.8
	R			83.1
I, III, V, VII	Spat.	.558	.300	46.4
	Num.	.511	.240	31.9
	D	.047	.060	14.5
	R			68.8
II, IV, V, VII	Spat.	.545	.285	45.1
	Num.	.526	.255	36.3
	D	.019	.030	8.8
	R			80.5
1, II, III, 4	Diff. Num.	.505	.303	42.6
	Easy Num.	.403	.171	16.7
	D	.102	.132	25.9
	R			39.2
II, IV, V, VI	Spat.	.548	.274	40.1
	Num.	.546	.272	39.7
	D	.002	.002	.4
	R			99.0
II, IV, 6	Spat.	.510	.290	38.7
	Num.	.511	.292	39.7
	D	−.001	−.002	−1.0
	R			102.6

TABLE XIV—(*Continued*)

Combinations of variables	*Averages*	K_1	K_2	$V_{(r)}$
3, 5, 7‡	Spat.	.572	.314	47.9
	Num.	.568	.310	46.2
	D	.004	.004	1.7
	R			96.5
4, 5, 7‡	Spat.	.511	.342	48.4
	Num.	.383	.186	18.2
	D	.128	.156	30.2
	R			37.6
3, V, VII‡	Spat.	.518	.298	48.5
	Num.	.475	.221	20.6
	D	.042	.077	27.9
	R			42.5
4, V, VII	Spat.	.507	.314	52.7
	Num.	.349	.100	8.6
	D	.158	.214	44.1
	R			16.3

* Difference between first average and second (generally, spatial minus numerical).

** Ratio between the second and the first average in percent.

‡ The variables were selected in this manner to emphasize the trends. Factor weights determined from the combinations of four variables containing the three variables involved are presented, but triad variances were determined from the three variables alone.

† Variables 2, 3 and II, IV were treated as 'strong' variables; variables 1, 4 and I, III as 'weak.'

consistent with one another although the tendency of the factor analyses, as has been pointed out before, is to decrease the variability of the weights within the first factor. Nevertheless, we may say that roughly variables 2 and 3 are the two strongest of the easy numerical set up and variables II and IV are the strongest of the difficult numerical variables. Within the spatial groups, variables 6 and VI tend to be the weakest. These weights and variances offer us an opportunity for some interesting combinations and comparisons which are presented in Table XIV.

If we consider first the easy numerical variables, we may combine the two easy pairs[28] with the easy and difficult spatials. Since the spatials were not well differentiated among themselves, we selected variables 5 and 7 and variables V and VII which tended to be the strongest in their particular set ups. In Table XIV, the results for each variable in its set up are presented in terms of the average weights for the first two components of the numerical and spatial variables, the average variance determined by the triad method, and the differences between these averages. In addition a ratio between the variances has been calculated.[29]

---

[28] For the purposes of this analysis, we are considering variables 2 and 3 as being 'strong' easy numerical variables; variables 1 and 4 as being 'weak' easy numerical variables; variables II and IV as being 'strong' difficult numerical variables; and variables I and III as being 'weak' difficult numerical variables.

We may see that, when the weak easy numerical variables and the strong easy ones are each compared with the spatial variables, the bond, by any measure in the table, is closer between the unlike variables when strong numericals are used than when weak ones are chosen. The same conditions hold when they are balanced against the difficult spatial variables. When the same types of set ups are introduced with the difficult numerical variables, the results obtained are comparable to those with the easy numerical ones. It is interesting to note that the absolute differences obtained here vary less than in the case of the easy numerical variables. When the strongest and the weakest easy numerical variables were combined with the spatial, the above tendencies were further emphasized, as may be seen from the table.[30]

This mode of approach enables us to show clearly that some of the numerical variables bear a closer relationship to some spatial ones than they do to other numerical variables. When the two weakest numerical variables (1, 4) are treated together with two relatively strong ones (II, III), the split between the weak and the strong appears to be very marked. When four variables matched as closely as possible for strength which has been determined from their relationships with their own type of variables, but two being numerical and two being spatial variables (II, IV, V, VI), are compared, we find the most remarkable correspondence between the two groups even though they differ in material. The differences for both the first and second components are only .002. The ratio between the triad variances is .99. Clearly, this shows that the general strength of a variable may transcend material bonds. A similar homogeneity is shown for variables II, IV, and 6.

---

[29] Since there was a close correspondence between the variances determined by the factor weights and by the triad method, it was thought unnecessary to transform the factor weights into measures which would permit a treatment by ratios.

[30] These facts are also brought out in comparing the cross correlations of the groups of numerical variables. For the strong easy numerical variables, the average cross correlation with the four spatial variables used here is .353 against, .264 for the weak. When the strongest and weakest variables alone are contrasted, the figures are .418 against .245. As one would expect, the difference is not so striking in the case of the difficult numerical variables, since it has been shown previously that the difficult variables as a group are strikingly homogeneous. The differences in cross correlation between the strong and the weak difficult numerical variables are consequently very small, the averages being .413 against .395. When the extreme variables are considered, the averages are .453 against .395.

c) The effect of level when material is constant.

The effect of level when material is constant may be studied from combinations of easy and difficult variables of the same material. These combinations may be subjected to a variety of treatments. In the case of both the numerical and spatial variables, a factor method and triad analyses may be used as was done in the preceding section. Because of the uniformity of the results obtained by the two methods of the previous analysis only triad analyses were used. In the case of the numerical variables, since they were originally derived from four tests, it is possible to set up tetrad difference patterns. This was done not only for all the easy and all the difficult variables taken as a unit, but for all the possible combinations of two easy with two difficult variables. The results of this analysis are presented in Table XV together with a tetrad analysis of variables II, IV, V, VI whose inclusion will be explained below.

TABLE XV

TETRAD ANALYSIS OF THE NUMERICAL VARIABLES TO SHOW EFFECT OF LEVEL

Variables	t_{abcd}	t_{abdc}	t_{acdb}	$t/P.E._{(t)}$*	*Chances***	*Ratio in % between the variances of easy and difficult variables*
1, 2, 3, 4	.0125	.0286 ± .0193	.0161	1.48	84	
I, II, III, IV	.0532 ± .0240	.0286	−.0246	2.22	93	
1, 2, III, IV	.0580 ± .0226	.0343	−.0237	2.57	96	60.1
1, 3, II, IV	−.0588 ± .0223	−.0187	.0401	2.64	96	69.2
1, 4, II, III	−.0134	.0271	.0405 ± .0186	2.18	93	39.2
2, 3, I, IV	−.0469	.0200	.0669 ± .0193	3.47	99	109.5
2, 4, I, III	.0035	.0235 ± .0177	.0200	1.33	82	84.1
3, 4, I, II	.0093	−.0143	−.0236 ± .0168	1.40	83	68.6
II, IV, V, VI	−.0039	.0021	.0060 ± .0237	.25	57	

* The probable error of the largest tetrad difference only in each combination of variables was calculated.

** Chances in 100 that the difference is greater than zero.

It is interesting to note that while no tetrad difference is as much as four times its probable error,[31] except for $t_{\text{II, IV, V, VI}}$, the values of the three differences in each group are not uniform and suggest the presence of group factors, that is, of special bonds between two or more variables. These bonds are in all probability not of the lowest possible order, as is shown by the fact that few of the differences conform to the simple patterns that Tryon (38) has shown first order group factors give rise to.[32] Using variables II, IV, V, VI, we obtain the smallest tetrad differences of the table, the ratio of the largest difference in this group to its probable error being but .25. This criterion again shows how these variables, though of different material, express a greater functional homogeneity than combinations of numerical variables which are alike in material.

When the average variance of the easy and difficult numerical variables are obtained from the above six set ups, we find, as one would expect from the correlations, that the variance of the difficult numerical variables is high. In four of the six set ups, this tendency is clearly shown by the ratios obtained. For variables 2, 3, I, IV, the easy variables show a slightly higher variance.

In Table XVI, is presented the average variance obtained for each variable when the determination is made using only variables of a different level together with the variables under consideration. For the numerical variables, the variances are in general higher for the difficult numerical than for the easy numerical variables. This merely emphasizes again that the difficult variables have higher intercorrelations with each other than the easy ones. In determining the triad variances, the numerator would always be composed of cross correlations, while the denominator for a difficult variable would always be an easy-easy correlation and vice versa for the denominator of an easy variable. Hence the results obtained.

In the case of the spatial variables, our results are most interesting. The numerator of our triad determination will again consist of easy-difficult correlations, while the denominator will be an easy-easy

---

[31] This is due largely to the small correlations of this study, the reasons for which we have gone into before. The small correlations do not drastically affect the probable error, but they do affect the tetrad difference. This results in making even the larger differences in our table appear to be 'insignificant,' that is, to be less than 4 P.E. greater than zero.

[32] For first order bonds between variables a and b or c and d in a typical tetrad equation, the pattern will be $+ + 0$, (A); for similar bonds between a and c or b and d, the pattern will be $- 0 +$, (B); and for the first order bonds between a and d or b and c, the pattern will be $0 - -$, (C). The only set of differences that suggests any one of these patterns is the one derived from variables 1, 3, II, IV which resembles the B type.

TABLE XVI

AVERAGE PERCENT VARIANCE FOR EACH VARIABLE WHEN EACH DETERMINATION IS MADE FROM TRIADS CONTAINING THE VARIABLE AND TWO OTHERS EACH OF THE SAME MATERIAL BUT OF DIFFERENT LEVEL

Variable	*Variance*	*Variable*	*Variance*
1	21.6	I	29.0
2	21.9	II	46.9
3	39.2	III	23.4
4	15.6	IV	47.7
Average	24.6	Average	36.8
5	17.6	V	39.0
6	26.5	VI	14.7
7	34.0	VII	29.0
Average	26.0	Average	27.6

Ratios in percent between easy and difficult numerical variances and between easy and difficult spatial variances were 66.8 and 94.2 respectively.

or a difficult-difficult correlation depending on the kind of variable in question. As a consequence, the variances will be lower (as the correlations from both the easy and difficult group are higher than the cross correlations). Between the two groups of variances, there was relatively no difference. When we compare the average variance of each group determined in the above manner with the variances determined for the same groups, but from variables of the same material and of the same level, the influence of level is strikingly illustrated. The average variance for the easy spatials has been dropped by our present treatment to 26.0% from the original average of 41.6%. For the difficult spatials, the drop is to 27.6% from an average of 42.0%. The respective ratios are: .63 and .66.

5. COMPARISON OF THE DERIVED VARIABLES WITH THE 'WHOLE' TEST VARIABLES

One of our major findings demonstrated that in both the numerical and in the spatial tests level of difficulty introduced a source of heterogeneity. In our spatial tests, while both the easy and difficult variables have almost exactly the same average intercorrelations in their own groups, the cross correlations are depressed. With regard to the numerical variables, because of the small value of the easy intercorrelations and the very much greater values of the difficult intercorrelations, the cross correlations are higher than those of the easy variables, but definitely lower than the intercorrelations of the difficult variables.

It becomes an interesting theoretical problem to discover what effect one could expect if the original correlations reported for the whole variables were determined on tests that happened to be more homogeneous with respect to the difficulty of the material they contained. A direct comparison of the original correlations for the whole variables would not tell the whole story because of the differences in size between the whole tests and the ones derived from them. Unless the function is of such a nature that its reliability is always equal to unity, we may expect that increased or decreased lengths of the test variables may affect correlation. The most direct method of comparison would be to double the size of our easy and difficult variables by adding to them items which are equivalent in difficulty to the original ones, or to take the original tests and delete half of the easy and half of the difficult items. Since neither of these methods was feasible in the present study, a theoretical equation was derived which would give the values of the original correlations if the tests were divided in half. One could think of a test variable as consisting of two equivalent halves and our problem would be to find the intercorrelations of the original tests when only one of these halves was used. We may think of an individual's deviational score on a given half-test as being represented by the formula (1) $x' - x/2$.[33] The product of two such deviations will be: (2) $(x' - x/2)\ (y' - y/2) = xy/4$.[34] The standard deviations may be obtained from (1) and will be equal to (3) $\sqrt{5 - 4r_{xx}'}/2$, and consequently the correlation between the two variables will be (4) $r_{xy}/\sqrt{5 - 4r_{xx}'}\ \sqrt{5 - 4r_{yy}'}$. It will be seen immediately that the validity of this formula depends on obtaining a true reliability. From our point of view, a true reliability is obtained through correlating two tests of the same function where the items in each test are equated for difficulty.[35] Since the original tests did not have a perfect order of difficulty, we may treat our results obtained from the use of formula (4) as only tentative. It will be noticed that with perfect reliabilities length will make no difference in correlations. This will be due simply to the fact that the addition of items could not under those circumstances alter the relative standings of individuals in the group with respect to each other.

---

[33] Only with perfect reliability would an individual's two 'half' scores be identical with each other. As discussed in the previous section, a trait may under some conditions be well measured without the reliability coefficient being equal to unity.

[34] We will have to assume that x' and x, and y' and y are close enough to each other that various combinations of "xy" products may be substituted for each other. Thus when 2 (2) is expanded we have $x'y' - x'y/2 - xy'/2 + xy/4 = xy/4$.

[35] In a sense, our newly derived measure does not differ functionally from the Spearman-Brown formula. In the Spearman-Brown formula, we are concerned with the relation of the length of a test to its 'self-correlation'; here with the effect of length on correlations with other variables. For the Spearman-Brown formula, Holzinger (17) finds a correspondence between the theoretically derived reliabilities and empirically determined ones up to a fourfold increase in the number of items.

In Table XVII are presented the original correlations among the whole variables, the original reliability coefficients and the correlations calculated from the above formula. It should be noted that the reliabilities for the spatial analogies and the spatial generalizations test are below the general run of the other reliabilities. This has the effect of slightly distorting the relationships shown by the original correlation table, particularly for the correlations between these two variables. Average correlations were calculated by first transmuting the correlation into Fisher's 'z.' In terms of Fisher's coefficient, the averages are, for the inter-numerical correlations, .318, for the inter-spatial, .328, and for the numerical-spatial, .300.

TABLE XVII

INTERCORRELATIONS OF WHOLE VARIABLE AND ESTIMATED INTERCORRELATIONS OF VARIABLES EQUAL IN SIZE TO ONE-HALF THE ORIGINAL ONES*

	1	*2*	*3*	*4*	*5*	*6*	*7*
1. Num. analog.	.8842	.4754	.4279	.4416	.4889	.3386	.3921
2. “ gen.	.3047	.8298	.4893	.4367	.4986	.4900	.3945
3. Arith. prob.	.2785	.2964	.8648	.5773	.6017	.4316	.4620
4. Num. series comp.	.3041	.2799	.3788	.8886	.5334	.4636	.5050
5. Spat. analog.	.2866	.2720	.3360	.3152	.7534	.4929	.5846
6. “ gen.	.1971	.2654	.2427	.2721	.2467	.7448	.4868
7. M.K. spat.	.3001	.2824	.3368	.2467	.3839	.3174	.9623

* Original correlations are above the diagonal. They were obtained from Smith (27, 24). Correlations for 'half' the variable are under the diagonal. Reliability coefficients obtained from Smith (p. 21) are in the diagonal.

These figures would indicate a considerable amount of homogeneity between the various groupings of correlations, a homogeneity which certainly is not indicated when the level factor is considered. Since our formula may have ironed out some of the differences existing between the groups of variables, we may consider the results obtained from analyzing the whole variable in regard to the relationships among the groups. The average numerical intercorrelation was .518 (Fisher's 'z'), the average for the spatial was .580[36] and the average cross group correlation was .504.[37] The ratio of the difference between the spatial and cross group correlations to its standard error is 1.69.

These averages are the results of various opposed tendencies as our earlier analysis has shown. Using the previous analysis as a base (See section 2 of this chapter), combinations of variables were made to bring out even more clearly the true relationships existing. Table XVIII presents the evaluation of the differences between various groups so selected. In this table, for the least reliable difference, the probability is 98 in 100 that the difference is greater than zero. By eliminating the intercorrelations among the easy numerical variables, we have remaining the three groups of variables which show that homogeneity of level and material make for higher correlations. These groups show in various combinations reliable differences with regard to groups that are heterogeneous either in their level or material characteristics or both. It is interesting to

TABLE XVIII

THE EFFECT OF LEVEL ON THE RELATIONSHIP BETWEEN VARIABLES OF THE SAME AND OF DIFFERENT MATERIAL*

Type of Combination	*nN + sS (M = .327) N = 18*	*Ns + NS (M = .374) N = 20*	*Ns + NS + ns + nS (M = .337) N = 40*	*NS (M = .371) N = 10*	*sS (M = .345) N = 6*
NN (M = .409)	.082		.072		
N = 6	2.93		3.00		
ss + SS (M = .442)		.078	.105		.097
N = 6		3.25	4.77		2.85
SS + NN (M = .420)				.049	
N = 9				2.04	
ss + SS + NN (M = .425)	.098	.051	.088		
N = 12	3.77	2.22	4.40		

* The first number in a cell represents the difference between the averages of its column and its row. For every pair compared, the average of the column is higher. The number below the difference is the ratio of the difference to its standard error. The lowest ratio of 2.04 implies that there are 98 chances in 100 of a difference greater than zero.

---

[36] With the correlation between variables 6 and 7 eliminated, this average drops to .536.

[37] With the cross correlations between numerical and spatial analogies and generalizations eliminated.

note that when the difficult numerical correlations are compared with groups that differ in material, the critical ratio is equal to 3.00. When the whole variable is used, despite the higher level of correlations, the difference is only .014 with a critical ratio of .34. Comparing all our homogeneous groups as a unit with all the cross group correlations, our critical ratio becomes 4.40. Table XVIII shows clearly how the influence of the easy numerical intercorrelations and the relative lack of homogeneity in function between easy spatial and difficult spatial variables has operated to give the results obtained when the whole variable is used.

Returning to our correlations for 'half' variables, it is interesting to note that the value .318 is identical with that obtained for the correlation among easy and difficult numerical variables and .328 is not much less than the .345 for the average correlation between easy and difficult spatial variables. Except for the easy numerical variables, the average correlation of any group homogeneous with respect to level and material is higher than our theoretical value for the whole variable correlations when the variables comprising the latter are treated as if they are equal in size to our easy and difficult variables. For the difficult numerical variables, the average correlation is .409, and for the easy spatials and difficult spatials, .440 and .443, respectively.

It would seem from these results, that granted a sufficiently high level of difficulty for a group, homogeneous tests (with regard to level) will correlate higher than tests that are heterogeneous. This seems to be borne out even when the material is different. The average correlation of the easy numerical and the easy spatial variables is .335 against .300 for our thoretical 'half' variables. On the difficult level, the value rises to .371.

This analysis based on theoretical coefficients is offered as a tentative affair, but it points the way to experiments on the effect of the uniformity of difficulty on the relationships among mental tests.

CHAPTER IV

CONCLUSIONS AND INTERPRETATIONS

The results seem to leave no doubt that the heterogeneity of the level of difficulty of test items introduces a true functional heterogeneity. Other factors being equal, differences in the level of the same kind of material result in a depression of the correlation between samples of it. This is most clearly shown in the case of the spatial variables where the general level of the correlations is equal for both levels taken separately, but where the correlations between variables differing in level is lower than either of the average values for the more homogeneous group. The easy and difficult items consequently cannot be thought of as measuring exactly the same ability. Whatever the psychological factors that lie behind the ability to work with spatial relationships may be, they were not homogeneous ones. When the relatively homogeneous functions that easy and difficult spatial items represent are combined and treated as if they represent a uniform function, we find that this heterogeneity is reflected in the correlational results obtained when they are compared with other variables. Though we were forced into using theoretical equations to obtain the necessary values for comparative purposes, it seems to be fairly clear that spatial items from different tests have more in common with each other when their difficulty level is fairly constant than when they are represented by the relatively diverse functions that items of the same material but of difficult level indicate.

In the case of the numerical variables, the effect of the level of difficulty is somewhat confused by the fact that the easy and difficult variables do not bear the same degree of relationship to their respective groups. The order of relationship is so low in the case of the easy numerical variables that they actually show a stronger bond with the difficult numerical variables than they do with each other. This bond is, however, considerably below that of the bonds to be found among the difficult numerical variables. These relationships indicate again a diversity of function which would make one hesitant to treat the numerical variables as if the function they were measuring were uniform. In comparing our correlations of the separate parts of these variables with correlations obtained from variables with mixed items, we again find it demonstrated that the diversity of function within a variable has marked effects on its relationships with the other functions. Our difficult numerical variables correspond more closely in function, if the correlation coefficient may be taken as an index of underlying functional relationships, with each other than do variables composed of mixed relationships.

The fact that diversity of difficulty may be an objective sign of diversity in function is reflected in the relationships uncovered between variables that differ in material, and by that criterion, one might offhand say, differ in underlying function. When the various tendencies that differences in level bring in are permitted to operate in amalgamated form, the conclusion is drawn that the relationships among numerical variables are no higher than the relationships between numerical variables and spatial ones. But with the effects of level teased out, it becomes apparent that this conclusion could be drawn only because of the following reasons: (1) Between easy numerical variables and easy spatials there exists a relationship greater in strength than among the easy numerical ones themselves. (2) This tends to be counterbalanced by the fact that the relationships between difficult numerical variables and easy and difficult spatials taken separately are smaller than those to be found among the difficult numerical variables. (3) The influence of the spatial variables were partly nullified by the fact that the whole variable was being expressed by functions which were not completely congruent with each other.

The existence of the diverse relationships which have contributed to the final influences that variables have on relations between themselves and other variables calls for some analysis of the implications of many of the facts that have been mentioned in our summary. When the constructor of a mental test devises a test to measure a particular ability, he is guided by common sense notions and experience with regard to what he is to include. There is no question but that the average test-maker can tell the difference between a numerical item and a verbal or spatial one and therefore he may feel a certain amount of justification in labeling a particular test a 'spatial' test and another one a 'numerical' test and so on. Items may be classified in various groups so as to make it possible for one to have a number of different tests in the same material category. It is implicit from the construction of the test that numerical tests, let us say, are measuring different forms of numerical functions. They may differ from each other, but they are more alike than some other variable defined in terms of its material. If we agree, however, that the magnitude of a correlation coefficient is a sign indicating that there is a certain amount of a common function between two variables which may or may not correspond to our formal classifications, we find that there are instances where the formal classifications are contradicted by the functional relationship that the correlation coefficient indicates. In such cases, we may think of the formal classification as playing the role of nothing more than an hypothesis which is to be tested, but which can be contradicted by relationships that are found to hold after investigation. This is true in our own case. Numerical problems are obviously different from spatial ones. Nevertheless, the average correlation among certain examples of the numerical function is .259, while it is .374 between other examples involving two apparently different functions, the numerical and the spatial.

If we are to consider our tests to be measuring an aspect of the function that psychologists commonly call "intelligence," we must be realistic in our approach and interpretations. We cannot see of what use in solving our problems in connection with the relationship between different forms of intelligent 'behavior' the usual definitions of intelligence can be. Thus, William Sterne tells us that intelligence involves conscious adaptive behavior to our environment (31). Spearman (29) develops three primary laws of cognition, to have them reduced to one by Wyatt (40). Terman and Colvin present definitions that seem to indicate hardly more than their special interests. Thus, to Terman, "an individual is intelligent in proportion as he is able to carry on abstract thinking." (32, 128.) Colvin (6, 136) finds that "an individual possesses intelligence in so far as he has learned, or can learn, to adjust himself to his environment." What we are really told in all these definitions, and others that may be introduced, is that there are modes of behavior that have been classified together into one category, intelligence. Whether the various functions indicated by these different forms of behavior are homogeneous is an experimental and a correlational problem.[38]

Consider from this point of view, our problem, restricted as it is only to mental tests, sheds some interesting light. For we find in fields far more narrow than most of those included by the definitions of intelligence that what we have considered to be definitions are in reality hypotheses. These hypotheses enabled us to set up an experiment, but the results had to be interpreted in a fashion contradictory to the hypotheses and new ones had to be set up. Since we feel that no matter how logical it may seem that certain characteristics should go together the fact that they actually need not do so, reduces the problem of the investigator to one of finding relationships. In the case of studies of this sort, the chief measure of relationship is the correlation coefficient. If that measure is assumed to indicate similar underlying functions we may then find that common sense hypothesis may be transcended by what this measure indicates.

---

[38] We do not wish to imply that these definitions do not serve an important role in linguistically organizing our problem. The concept of intelligence may be broad enough to include various functions with different degrees of relationship to each other and yet be of value for psychological analysis. Our point is that no definition will tell us anything about the extent of these relations. Finding the latter are experimental problems. By suggesting a certain uniformity in intelligent behavior, these definitions often have the effect of making us forget that such diverse relationships may exist.

In our own study, such findings have been reported. There is no need of reviewing all the instances in which true relationships have superseded formal classifications. The real problem is to find the conditions under which this occurs. Wherever possible this has been indicated in the text.

In regard to the construction of mental tests, the results obtained seem to have some bearing. It seems clear, that provided that the items are not too easy, tests of the same material will correlate higher with each other when they are of about the same level of difficulty. It is possible that if tests are too difficult correlations will drop. Since we did not exhaust the difficulty range for our subjects, we can say nothing in regard to this point on the basis of our results.

Our results, however, do not give us any criteria enabling us to say with certainty how easy a test must be for its low difficulty value to overcome the influence of homogeneity of level. That it will vary with the group under consideration is most probable. That it will vary with the kind of material used is also likely. Therefore, for any test constructed, an analysis of various combinations of items is in order before one can definitely say at which levels of difficulty the effects even of homogeneity of difficulty may be overcome.

BIBLIOGRAPHY

1. Alexander, W. P. Intelligence; concrete and abstract. Brit. J. Psych. Mon. Suppl., 1935, XIX.
2. Anastasi, A. A group factor in immediate memory. Arch. Psych., No. 120, 1930.
3. Asch, S. E. A study of change in mental organization. Arch. Psych., No. 195, 1936.
4. Brolyer, C. R. A formula for the mean tetrad. J. Gen. Psych., 1932, 6, 212–214.
5. Clark, J. R. The relation of speed, range, and level to scores in intelligence tests. New York: Teachers College, Columbia University, 1925.

6. Colvin, S. S. Intelligence and its measurement: A symposium. J. Ed.. Psych., 1921, 12, 136–139.
7. DuBois, P. H. A speed factor in mental tests. Arch. Psych. No. 141, 1932.
8. Dunlap, J. W., and Kurtz, A. K. Handbook of statistical nomographs, tables and formulas. Yonkers (N. Y.): World Book Co., 1932.
9. Fisher, R. A. Statistical methods for research workers. London: Oliver, 1932 (4th ed.).
10. Freeman, F. S. The factors of speed and power in tests of intelligence. J. Exp. Psych., 1931, 14, 83–90.
11. Furfey, P. H., and Daly, T. F. Product moment correlations as a research technique in education. J. Ed. Psych., 1935, 26, 206–211.
12. Garrett, H. E. Statistics in psychology and education. New York: Longmans, 1929.
13. ————. A study of the CAVD intelligence examination. J. Ed. Res., 1930, 21, 103–108.
14. ————, and Anastasi, A. The tetrad-difference criterion and the measurement of mental traits. Ann. N. Y. Acad. Sci., 1932, 33, 235–281.
15. ————, Bryan, A. I., and Perl, R. E. The age factor in mental organization. Arch. Psych., No. 176, 1935.
16. Hamid, S. A. Some factors of effectiveness in mental ("intelligence") tests. Brit. J. Psych., 1925–1926, 16, 100–115.
17. Holzinger, K. J. Note on the use of Spearman's prophecy formula for reliability. J. Ed. Psych., 1923, 14, 302–305.
18. ————. Statistical resumé of the Spearman two-factor theory. Chicago: University of Chicago Press, 1930.
19. Hotelling, H. Analysis of a complex of statistical variables into principal components. J. Ed. Psych., 1933, 24, 417–441; 498–520.
20. Kelley, T. L. Statistical method. New York: Macmillan, 1924.
21. ————. Crossroads in the mind of man. Palo Alto (Calif.): Stan. U. Press, 1928.
22. ————. Essential traits of mental life. Cambridge: Harvard Univ. Press, 1935.
23. Peatman, J. G. A study of factors measured by the Thorndike Intelligence Examination for high school graduates. Arch. Psych., No. 128, 1931.
24. ————, and Greenspan, I. The reliability of a questionnaire on superstitious beliefs of elementary school children. J. Abn. & Soc., 1935–1936, 30, 208–221.
25. Schiller, B. Verbal, numerical and spatial abilities of young children. Arch. Psych., No. 161, 1934.
26. Schneck, M. R. The measurement of verbal and numerical abilities. Arch. Psych., No. 107, 1929.
27. Smith, G. M. Group factors in mental tests similar in material or in structure. Arch. Psych., No. 156, 1933.
28. Spearman, C. Correlations of sums and differences. Brit. J. Psych., 1912–1913, 5, 417–426.
29. ————. The nature of intelligence and the principles of cognition. London: Macmillan and Co., 1923.
30. ————. The abilities of man. New York: Macmillan, 1927.
31. Stern, W. The psychological methods of testing intelligence. Baltimore: Warwick and York, 1914.
32. Terman, L. M. Intelligence and its measurement: A symposium. J. Ed. Psych., 1921, 12, 127–133.
33. Thorndike, E. L., *et al.* The measurement of intelligence. New York: Teachers College, Columbia University, 1926.
34. Thurstone, L. L. Reliability and validity of tests. Ann Arbor (Mich.): Edwards Bros., Inc., 1931.
35. ————. A simplified Multiple Factor method and on outline of computations. Chicago: University of Chicago Press, 1933.
36. ————. The vectors of mind. Chicago: University of Chicago Press, 1935.